Spine

SPINE Volume 45, Number 23, pp 1649–1660

EPIDEMIOLOGY

OPEN

Trajectories of Disability and Low Back Pain Impact

2-year Follow-up of the Groningen Spine Cohort

Alisa L. Dutmer, MSc,[a] Henrica R. Schiphorst Preuper, MD, PhD,[a,b] Roy E. Stewart, PhD,[c] Remko Soer, PhD,[b,d] Michiel F. Reneman, PhD,[a] and André P. Wolff, MD, PhD[b,e]

Study Design. Prospective cohort study.

Objective. The aim of this study was to identify treatment response trajectories in patients with low back pain (LBP) during and after multidisciplinary care in a tertiary spine center, and to examine baseline patient characteristics that can distinguish trajectories.

Summary of Background Data. Treatment response is often heterogeneous between patients with LBP. Knowledge on key characteristics that are associated with courses of disability could identify patients at risk for less favorable outcome. This knowledge will help improve shared decision-making.

Methods. Adult patients with LBP completed questionnaires on disability (Pain Disability Index) and LBP impact (Impact Stratification of the National Institutes of Health minimal dataset) at baseline, 6, 12, 18, and 24 months' follow-up. Latent class analyses were applied to identify trajectories of disability and LBP impact. Baseline sociodemographic and clinical patient characteristics were compared between trajectory subgroups.

Results. Follow-up was available for 996 patients on disability and 707 patients on LBP impact. Six trajectories were identified for both outcome measures. Three disability trajectories remained stable at distinct levels of severity (68% of patients) and three trajectories showed patterns of recovery (32%). For LBP impact there was one stable trajectory (17%), two slightly improving (59%), two recovering (15%), and one with a pattern of recovery and relapse (15%). Significant differences between trajectories were observed for almost all baseline patient characteristics.

Conclusion. On average, patients show moderate improvements in disability and LBP impact 2 years after visiting a multidisciplinary tertiary spine center. However, latent class analyses revealed that most patients belong to subgroups experiencing stable levels of disability and LBP impact. Differences in baseline patient characteristics were mostly associated with baseline levels of functioning, instead of (un)favorable outcome during follow-up.

Key words: chronic pain, cohort study, disability, follow-up study, functional limitation, latent class analysis, low back pain, minimal clinically important difference, multidisciplinary care, recovery, responder analysis, trajectories.

Level of Evidence: 2

Spine 2020;45:1649–1660

Many efforts have been undertaken to reduce the burden of LBP, but unfortunately there is only low to moderate level of evidence for the effectiveness of LBP interventions in the short, medium, and long term.[1–4] Clinical trial results are often interpreted based on summary measures of treatment effects. However, when focusing on the results of individual patients it is apparent that treatment response is often heterogeneous.[5] There are subgroups of patients that achieve successful outcomes (responders) and subgroups that achieve little to no improvements, or even worsen (nonresponders). Although there has been a change of research-perspective toward the individual patient, the inclusion of responder analyses in chronic LBP trials is still rare.[5,6]

Heterogeneous treatment response is not limited to outcome differences on an individual level at a single time-point, but it can also manifest as different trajectories of pain or disability during and after treatment. Previous studies on LBP trajectories identified several patterns of pain post-treatment:

From the [a]University of Groningen, University Medical Center Groningen, Department of Rehabilitation Medicine, Groningen, The Netherlands; [b]University of Groningen, University Medical Center Groningen, Pain Center, The Netherlands; [c]University of Groningen, University Medical Center Groningen, Department of Public Health, Groningen, The Netherlands; [d]Saxion University of Applied Sciences, Expertise Center of Health and Movement, Enschede, The Netherlands; and [e]University of Groningen, University Medical Center Groningen, Department of Anaesthesiology, Groningen, The Netherlands.

Acknowledgment date: February 28, 2020. First revision date: May 7, 2020. Acceptance date: June 16, 2020.

The manuscript submitted does not contain information about medical device(s)/drug(s)

Healthy Ageing Pilots UMCG (process nr. CDO17.0038/2017-2/331) funds were received in support of this work.

No relevant financial activities outside the submitted work.

Address correspondence and reprint requests to Michiel F. Reneman, PhD, Center for Rehabilitation, P.O. Box 30002, 9750 RA Haren, The Netherlands; E-mail: m.f.reneman@umcg.nl

DOI: 10.1097/BRS.0000000000003647

recovery, persistent severe LBP, neither recovery nor persistent severe LBP, and fluctuating LBP.[7–9] Although patient expectations of treatment are (initially) focused primarily on pain reduction,[10] health care professionals mainly aim for functional improvement and increased quality of life. We identified two studies that reported trajectories of levels of disability in patients with back pain.[11,12] In a sample of elderly patients seeking primary care, 6% showed functional recovery after 12 months' follow-up, whereas others remained at stable levels of disability. A shorter pain duration, higher confidence in improvement pretreatment, and less comorbidities were associated with improvement in LBP-related disability.[11] Furthermore, approximately one-third of patients undergoing surgery for spinal stenosis experienced rapid improvements in disability, whereas the other patients belonged to fair or poor outcome trajectory groups.[12]

Knowledge on key characteristics that are associated with varying courses of disability could help identify clinically relevant subgroups of patients who are likely to (not) respond to LBP treatment and may improve shared decision-making. This study is part of the Groningen Spine Cohort (GSC), an ongoing 10-year prospective cohort study of adult patients with LBP visiting a multidisciplinary tertiary care spine center.[13] The two main objectives of the present study are: to identify trajectories of disability and LBP impact in patients with LBP over a period of 24 months, during and after treatment, and to examine whether the patients in these trajectories can be characterized by sociodemographic and clinical patient characteristics. Furthermore, we will examine how trajectories relate to the number of individual patients that reach a minimal clinically important outcome (responders) during follow-up.

MATERIALS AND METHODS

Study Design

Patients that participate in the GSC (n = 1502) were included in this study.[13] Baseline and follow-up data were acquired between July 2015 and July 2019 through digital questionnaires and medical records. Measurements took place at baseline and at 6, 12, 18, and 24 months' follow-up. The Medical Ethical Committee of the University Medical Center Groningen, the Netherlands, provided a waiver (M15.169472) for the data collection of the GSC with respect to medical ethical permission. Handling of the data was done in accordance with the guidelines for Good Research Practice.[14] The STROBE statement on cohort studies was used as a reporting guideline.[15]

Patients and Setting

Patients were referred to the Groningen Spine Center, a tertiary multidisciplinary center of a university hospital in the north of the Netherlands by their general practitioner or a medical specialist. Inclusion criteria for the GSC were LBP and/or leg pain and an age between 18 and 65 years' old. For the present study, patients with at least one follow-up measurement on one of the primary outcomes were included. Patients that did not understand Dutch language or had no Internet access were excluded. All patients were informed on the purpose of the study and signed an informed consent. Treatment at the university hospital was care as usual and could consist of minimal intervention only (pain education, information, and reassurance), or could be combined with multidisciplinary rehabilitation, surgery, pain anesthesiology treatment, referral to primary care, and/or other if needed.

Measures

Primary Outcomes

Disability

The Pain Disability Index (PDI) measures self-reported pain interference for seven categories of daily life activities: family/home responsibilities, recreation, social activity, occupation, sexual behavior, self-care, and life support activity.[16] Each item is constructed on an 11-point numeric rating scale with 0 indicating no disability and 10 indicating maximum disability. Total scores range from 0 to 70 where higher scores represent greater disability due to pain. The Dutch version of the PDI is responsive to change and 2-week test-retest reliability is good.[17,18] The PDI was administered at all measurements.

LBP Impact

LBP impact was measured with the Impact Stratification of the National Institutes of Health (NIH) minimal dataset for chronic LBP.[19] The Impact Stratification consists of nine items: a numeric rating scale (NRS) of pain intensity (range 0–10), four items on pain interference (range 1–5), and four items on physical function (range 1–5). The total score ranges from 8 (least impact) to 50 (most impact). The NIH Research Task Force classified scores as mild (8–27 points), moderate (28–34 points), or severe impact (≥35 points).[19] The Impact Stratification has moderate to strong correlation with concurrent measures[19] and is responsive to change.[20] The Impact Stratification was administered at baseline, 12 months', and 24 months' follow-up.

Secondary Outcomes

The remaining items of the NIH minimal dataset were administered at baseline to collect data on demographics, smoking status, duration of pain, pain intensity, treatment(s) used for LBP, and depression and catastrophizing.[19] Four items on depression and two items on catastrophizing formed a factor ''depression and catastrophizing' (range 0–100, higher scores represent more feelings of depression and catastrophizing).[20,21] The Euroqol-5D (EQ5D), a 5-item questionnaire, was used to measure quality of life.[22] For scoring the EQ5D, the Dutch utility index was used with values ranging from −0.33 to 1.00, with higher values reflecting higher quality of life.[23] The Work Ability Score (WAS) measured self-reported current work ability compared to lifetime best (range 0–10).[24] Additionally, patients

were asked about their work status and whether they (applied for) receiving disability benefits.

At the first follow-up measurement following their discharge, patients were asked on a seven-point Likert scale if they were satisfied with the care they received at the Spine Center. The number of times patients responded with "very satisfied" or "satisfied" was reported in the results. Information on treatment(s) that patients received at the university hospital during follow-up were obtained from electronic patient records. The following treatment groups were reported: multidisciplinary rehabilitation, pain anesthesiology treatment, surgery, other, a combination, or "no further treatment at the university hospital" (other than pain education, and information).

Analysis

Missing Data

We checked the influence of any missing data by comparing the baseline characteristics (PDI and Impact Stratification scores, pain intensity, gender, age, education level, and duration of LBP) of patients with a complete follow-up to that of patients with incomplete follow-up.

Trajectories of Disability and LBP Impact

Latent class analyses (LCA; Mplus version 7.1) were performed to identify subgroups of patients with different trajectories of disability and LBP impact over time. Full information maximum likelihood (FIML) was used to take into account missing data. FIML does not impute missing data, but estimates parameters directly using all the information that is already contained in the incomplete data set. We tested models that ranged from two to nine classes. The Bayesian Information Criterion (BIC), entropy, class sizes, and interpretability of competing solutions were used to determine the number of classes that best fitted the data. A lower BIC indicates a better fit and a entropy reflects the quality of classification.[25,26] Trajectories were classified by the level of severity at baseline (*e.g.*, low, moderate, high) and its course over time (*e.g.*, stable, improving, recovery).

Patient Characteristics

Chi-square tests and Kruskal-Wallis tests or analysis of variances, depending on the distribution of the data, were used to examine whether patient characteristics differed between disability and LBP impact trajectories resulting from the LCA. Post hoc testing was performed to determine which trajectories were significantly different from all other trajectories (for continuous variables) or differed from the expected value (ordinal values). A P value of <0.05 was considered statistically significant.

Responder Analyses

Minimal clinically important change (MCIC) was used to determine whether a patient was a responder on disability and LBP impact at follow-up. We determined the MCIC for the PDI in a separate analysis (see text and Tables, Supplemental Digital Content 1, http://links.lww.com/BRS/B609, for details on methodology and results) in a subset of patients (n = 274) of the GSC. The optimal cut-off for the MCIC (with the highest sensitivity and specificity) was determined at an improvement of 44% in PDI score from baseline. MCICs for the Impact Stratification are also baseline-dependent. Patients with mild LBP impact at baseline (8–27 points) had to improve 7.5 points and patients with moderate (28–34 points) and severe (≥35 points) impact had to improve respectively 11.5 and 12.5 points to be considered a responder.[20] For the PDI and Impact Stratification the proportion of responders per follow-up measurement was reported for the total sample and for each individual latent class group.

RESULTS

A minimum of one follow-up measurement was available for 996 patients on the PDI and 707 patients on Impact Stratification. Baseline characteristics are presented in Table 1.

Complete follow-up (6, 12, 18, and 24 months) was available for 248 of 996 patients (25%) on the PDI and 296 of 707 patients (42%) on Impact Stratification. Patients with complete follow-up were on average 5 years older than patients with incomplete follow-up (PDI sample: 51.4 ± 11.3 years *vs.* 46.7 ± 13.0, $P < 0.01$; Impact Stratification sample: 51.1 ± 11.6 years *vs.* 47.4 ± 12.8, $P < 0.01$).

Overall, mean scores on PDI and Impact Stratification improved significantly ($P < 0.01$) on all measurements compared to baseline (Table 2). Correlations between PDI and Impact Stratification were $r = 0.76$ at baseline, $r = 0.88$ at 12 months, and $r = 0.87$ at 24 months' follow-up (all $P < 0.01$).

Model Selection

The BIC suggested that for both the PDI and Impact Stratification a six-class model was the best fit (see Table, Supplemental Digital Content 2, http://links.lww.com/BRS/B610, for model selection criteria results). The models ranging from four to seven classes did not differ in the ability to separate classes, with entropy ranging between 0.67 and 0.71. The smallest class in the six-class model for Impact Stratification contained only 37 patients (5%), but in comparing the trajectories of the five-class and six-class model the addition of an extra class was judged to be conceptually meaningful. The six-class model introduced a trajectory (10%) where patients show a relapse in LBP impact.

Trajectories of Disability and LBP Impact

Three of six disability trajectories (68% of patients) remained relatively stable over 2 years at distinctly different levels of severity (class 2 "Moderate Disability, Stable," class 5 "Moderate High Disability, Stable," and class 6 "High Disability, Stable"; Figure 1). Two trajectories started at moderate to high disability and recovered fast (class 1 "Moderate High Disability, Fast Recovery") or with a delay (class 4 "Moderate High

TABLE 1. Baseline Characteristics for PDI and Impact Stratification Samples

Characteristic	PDI Sample (n = 996)	Impact Stratification Sample (n = 707)
Age, y, mean (SD)	47.9 (12.8)	49.0 (12.4)
Sex (female), n (%)	576 (58)	400 (57)
Current smoker, n (%)	298 (30)	186 (26)
Education level, n (%)		
Low	349 (35)	245 (35)
Middle	345 (35)	240 (34)
High	231 (23)	170 (24)
Other/unknown	71 (7)	52 (7)
Work status, n (%)		
Working	265 (27)	191 (27)
Partial sick leave	179 (18)	122 (17)
Sick leave	166 (17)	115 (16)
Unemployed	386 (39)	279 (40)
Duration LBP, n (%)		
<3 mo	23 (2)	15 (2)
3 mo–1 y	150 (15)	94 (13)
1 y–5 y	351 (35)	253 (34)
>5 y	472 (47)	345 (49)
Previous treatment(s) for LBP, n (%)		
Low back surgery	265 (27)	193 (27)
Opioids	537 (54)	384 (54)
Injections	231 (23)	167 (24)
Exercise therapy	882 (89)	628 (89)
Psychological counseling	155 (16)	109 (15)
Pain intensity (0–10), mean (SD)	6.7 (1.7)	6.7 (1.6)
PDI (0–70), mean (SD)	37.7 (13.9)	37.7 (13.6)
Impact Stratification (8–50), mean (SD)	35.1 (7.3)	35.1 (7.2)
EQ5D utility score (−0.33 to 1.00), median (IQR)	0.56 (0.17;0.73)	0.57 (0.19;0.73)
NIH minimal dataset: depression and catastrophizing (0–100), median (IQR)	33.3 (16.7;58.3)	33.3 (16.7;58.3)
Work ability score (0–10), mean (SD)	3.9 (2.9)	3.8 (2.9)
Receives (or applied for) disability benefits due to LBP, n (%)	244 (25)	169 (24)

IQR indicates interquartile range; LBP, low back pain; n, number of patients; mo, months; PDI, Pain Disability Index; SD, standard deviation; y, years.

Disability, Late Recovery"), and one trajectory started at low to moderate disability and gradually recovered during follow-up (class 3 "Low Moderate Disability, Slow Recovery").

For LBP impact, there was one stable trajectory of severe impact (class 4 "Severe Impact, Stable"), and two relatively stable trajectories that improved from severe impact to moderate (class 2 "Severe Impact Improving to Moderate") and moderate impact to mild (class 6 "Moderate Impact Improving to Mild") (Figure 2). The remaining three trajectories showed recovery after one-year follow-up; however,

TABLE 2. Mean Disability and LBP Impact Scores at Baseline and Follow-up

Outcome	Baseline	6 mo	12 mo	18 mo	24 mo
PDI (0–70)					
n, (%)	996 (100)	773 (78)	663 (64)	566 (57)	376 (38)
Mean (SD)	37.7 (13.9)	32.0 (15.7)*	28.3 (17.0)*	28.8 (16.7)*	27.8 (17.2)*
Impact stratification (8–50)					
n, (%)	707 (100)	n/a	629 (89)	n/a	374 (53)
Mean (SD)	35.1 (7.2)	n/a	28.2 (10.2)*	n/a	28.3 (10.2)*

IQR indicates interquartile range; LBP, low back pain; mo, months; n, number of patients; n/a, not assessed; PDI, Pain Disability Index; SD, standard deviation.

**Significant improvement compared to baseline, $P < 0.01$.*

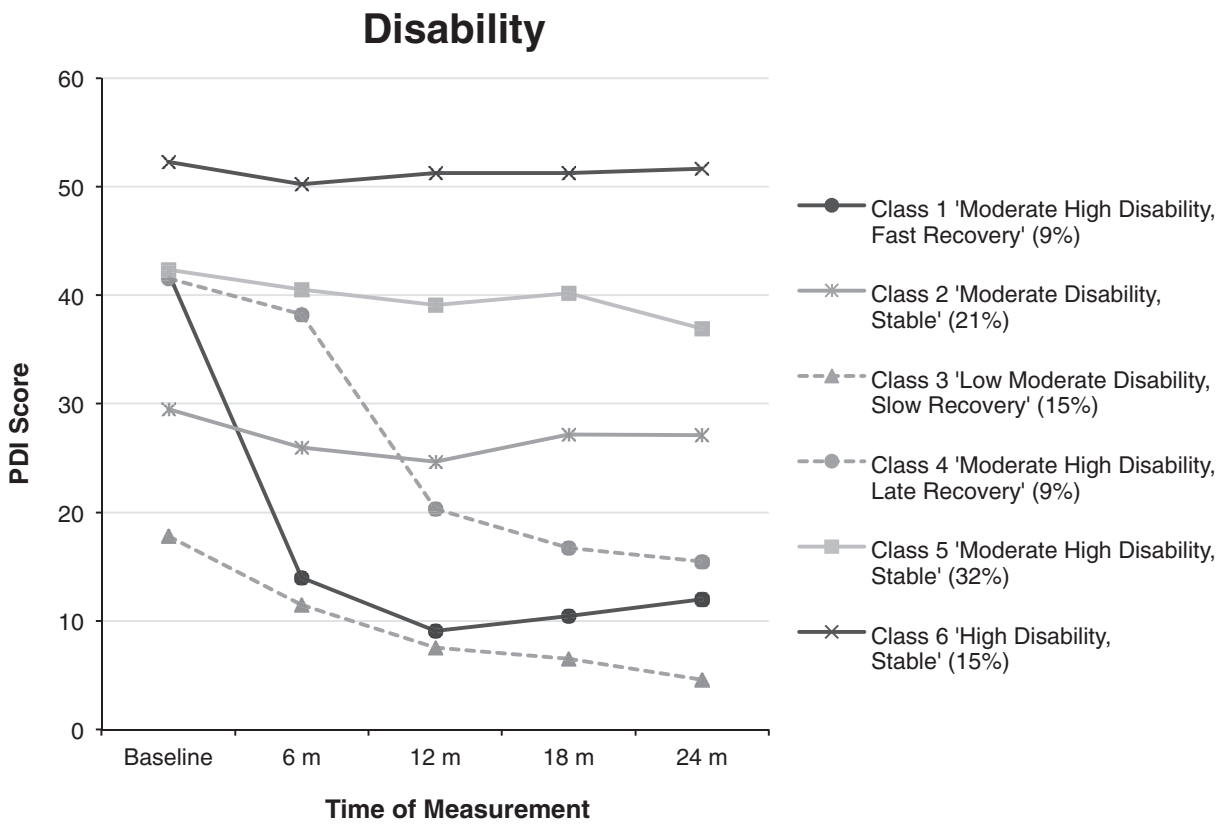

Figure 1. Disability trajectories per latent class group. PDI indicates Pain Disability Index.

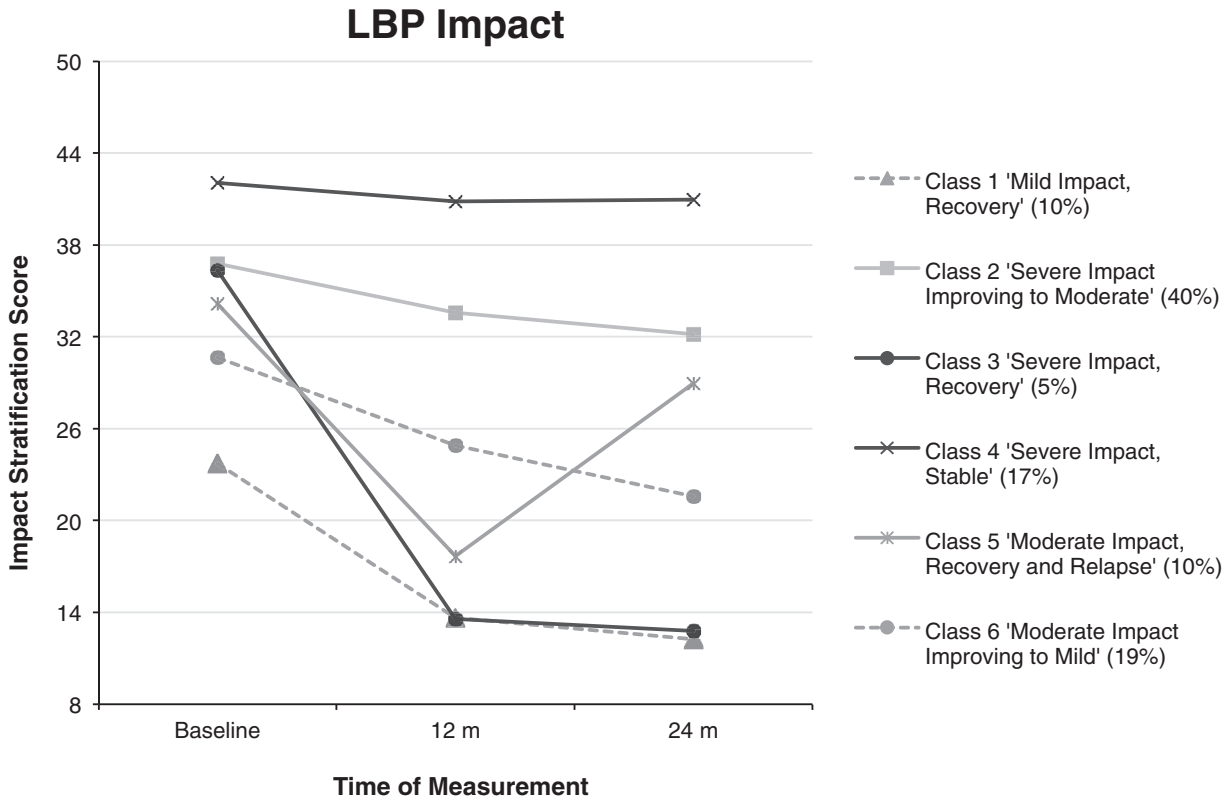

Figure 2. LBP Impact trajectories per latent class group. LBP indicates low back pain.

one group (class 5 "Moderate Impact, Recovery and Relapse", 10% of patients) relapsed to moderate LBP impact at the 2-year measurement.

Association of Patient Characteristics With Disability and LBP Impact Trajectories

Except for sex and having tried exercise therapy as a treatment in the past, there were differences on all other baseline patient characteristics across disability and LBP impact trajectories (Tables 3 and 4). The Low Moderate Disability, Slow Recovery trajectory group (class 3) was higher educated, more often worked fully, contained less smokers, has had less previous treatment(s) for LBP, and scored better on measures of pain, functioning, and quality of life at baseline compared to other trajectories (Table 3). These patients were also most often satisfied with the care that they received at the spine center. For the High Disability, Stable trajectory group (class 6) it was the other way around. The patients from trajectories starting at moderate or moderate to high disability scored in-between on most measures. During follow-up, the Moderate High Disability, Late Recovery trajectory group (class 4) more often had surgery and were more likely to have had at least one other kind of treatment besides education and advice, compared to other trajectories.

A similar pattern was observed for LBP impact, where the Mild Impact, Recovery group (class 1) was higher educated, more often worked fully, had a shorter duration of LBP, tried less previous treatment(s) for LBP, and scored better on measures of pain, functioning, and quality of life compared to other trajectories (Table 4). The Severe Impact, Stable trajectory group (class 3) showed mostly opposite results and also contained the highest proportion of smokers. All other trajectories scored in-between on baseline patient characteristics. The Moderate Impact, Recovery and Relapse trajectory group (class 5) more often underwent multidisciplinary rehabilitation during follow-up compared to other trajectories.

Responders

For disability, >75% of patients in each recovery group were responders (>44% improvement from baseline) at the final follow-up (Table 5). There were <15% responders in the stable groups.

The proportion responders was also >75% in the recovery groups of the LBP impact trajectories. The recovery and relapse group (class 5) went from 87% responders at 6 months back to 22% responders at 12 months' follow-up. Class 6, Moderate Impact Improving to Mild, improved on average nine points from baseline to 2-year follow-up and contained 42% responders.

DISCUSSION

In this study we identified treatment response trajectories in patients with LBP during and after multidisciplinary care in a tertiary spine center. Latent class analyses revealed six trajectories for both disability and LBP impact. Although almost one-third of patients experience substantial improvements, most patients remain stable at distinct levels of severity during 2-year follow-up. Significant differences between trajectories were observed for almost all baseline patient characteristics. However, patient characteristics were mostly associated with baseline levels of functioning and did not distinguish recovery trajectories from stable trajectories.

Two main patterns were distinguished for disability trajectories: recovery or persistent disability. Similar results were found in a study in elderly patients with LBP.[11] The number of patients belonging to a recovery trajectory group was larger in our sample (33% *vs.* 6%), which may in part be attributed to an age difference (average: 48 ± 13 *vs.* 74 ± 7 years) in study samples. Stable trajectories of disability were observed at varying levels of severity, with most patients (32%) experiencing moderate to high disability. Patients consulting secondary LBP care are more likely to experience persistent pain than those in primary care or in the general population.[27] It is not unexpected that patients presenting in tertiary care are more likely to experience highly persistent levels of pain or disability. Half of the patients in our study have experienced LBP for >5 years at baseline.

Although there were high correlations between scores on the PDI and Impact Stratification, LBP impact trajectories varied slightly more than disability trajectories. Also, two of three (somewhat) stable impact trajectory groups still experienced minor improvement. A study in patients undergoing spinal surgery similarly found that trajectories of disability were slightly more stable than trajectories of pain.[12] They suggested that pain (pain intensity is also one of the items of the Impact Stratification) may be more modifiable than disability following spinal surgery. The duration of follow-up also allowed us to identify a group of patients that relapsed on LBP impact after hitting the 1-year mark. Recurrence of LBP is very common within 12 months after recovery.[28] The number of responders in this patient group dropped from 87% to 22%. This further illustrates that a single time-point measure of treatment success can be misleading. Nonetheless, there were no characteristics at baseline that could help differentiate patients with a relapse in LBP impact from those without.

Differences in baseline patient characteristics were mostly associated with baseline levels of functioning. However, we observed a trend that some characteristics, most of which were previously reported in studies on predictors of LBP treatment success, might be related to functional recovery. These were a younger age,[11,29] being a nonsmoker,[11,30] and having a shorter duration of LBP.[31,32] Furthermore, severely impacted patients seemed to have a better chance at recovery when at sick leave but employed, than when unemployed. Characteristics that we did not measure, but might also be useful to assess at baseline are fear avoidance,[32,33] recovery expectation,[11,34,35] and medical comorbidity.[11] Finally, sex,[11,36] pain intensity,[29,36] level of disability[29,32,37] education level,[30,34,35] receiving disability benefits,[36] and mood/distress[29,32,33,38] have also been

TABLE 3. Characteristics of Patients With Different Disability Trajectories.

Characteristic	Class 1: Moderate High Disability, Fast Recovery	Class 2: Moderate Disability, Stable	Class 3: Low Moderate Disability, Slow Recovery	Class 4: Moderate High Disability, Late Recovery	Class 5: Moderate High Disability, Stable	Class 6: High Disability, Stable	ANOVA/ Kruskal-Wallis/χ^2
N (%)	87 (9)	206 (21)	148 (15)	93 (9)	314 (32)	148 (15)	
Baseline							
Age, y, mean (SD)	45.9 (10.6)	49.5 (12.5)	44.7 (13.1)	48.8 (13.6)	48.3 (13.1)	48.5 (12.1)	p = 0.01
Sex, n (%) female	52 (60)	115 (56)	78 (53)	55 (59)	193 (62)	83 (56)	ns
Current smoker, n (%)	22 (25)	53 (26)	20 (14)*	24 (26)	106 (34)	73 (49)*	p = 0.01
Education level, n (%)							p = 0.01
Low	32 (37)	68 (33)	30 (20)*	28 (30)	121 (39)	70 (47)*	
Middle	31 (36)	66 (32)	48 (32)	36 (39)	112 (36)	52 (35)	
High	16 (18)	58 (28)	63 (43)*	23 (25)	55 (18)	16 (11)*	
Other/unknown	8 (9)	14 (7)	7 (5)	6 (7)	26 (8)	10 (7)	
Work Status, n (%)							p = 0.01
Working	21 (24)	70 (34)	84 (57)*	25 (27)	51 (16)*	14 (10)*	
Partial sick leave	21 (24)	52 (25)	28 (19)	15 (16)	43 (14)	20 (14)	
Sick leave	25 (29)*	21 (10)	6 (4)*	18 (19)	63 (20)	33 (22)	
Unemployed	20 (23)*	63 (31)	30 (20)*	35 (38)	157 (50)*	81 (55)*	
Duration LBP, n (%)							p = 0.02
<3 mo	3 (3)	2 (1)	0 (0)	2 (2)	11 (4)	5 (3)	
3 mo–1 y	18 (21)	30 (15)	24 (16)	13 (14)	49 (16)	16 (11)	
1 y–5 y	32 (37)	81 (39)	62 (42)	40 (43)	89 (28)*	47 (32)	
>5 y	34 (39)	93 (45)	62 (42)	38 (41)	165 (53)	80 (54)	
Previous treatment(s) for LBP, n (%)							
Low back surgery	25 (27)	35 (17)*	18 (12)*	23 (25)	94 (30)	70 (47)*	p < 0.01
Opioids	47 (54)	98 (48)	36 (24)*	53 (57)	194 (62)*	109 (74)*	p < 0.01
Injections	17 (20)	42 (20)	15 (10)*	20 (22)	72 (23)	65 (44)*	p < 0.01
Exercise therapy	76 (87)	178 (86)	132 (89)	86 (93)	277 (88)	133 (90)	ns
Psychological counseling	9 (10)	27 (13)	2 (1)*	14 (15)	58 (19)	45 (30)*	p < 0.01
Pain intensity (0–10), mean (SD)	6.6 (1.5)	6.3 (1.6)	5.6 (1.8)*	6.9 (1.5)	7.1 (1.4)	7.8 (1.3)*	p < 0.01
PDI (0–70), mean (SD)	44.0 (8.1)	28.7 (8.3)*	17.6 (7.9)*	43.1 (8.5)	42.2 (8.7)	53.6 (6.7)*	p < 0.01
Impact stratification (8–50), mean (SD)	35.9 (5.5)	32.3 (5.8)*	26.1 (6.8)*	36.7 (5.7)	37.1 (4.8)	42.5 (4.2)*	p < 0.01
EQ5D utility score (−0.33 to 1.00), median (IQR)	0.35 (0.19;0.65)	0.69 (0.38;0.78)*	0.78 (0.69;0.81)*	0.30 (0.17;0.65)	0.30 (0.17;0.65)	0.17 (0.07;0.27)*	p < 0.01
NIH minimal dataset: depression & catastrophizing (0–100), median (IQR)	33.3 (25;50)	25.0 (12.5;42.7)	14.9 (0;20.8)*	41.3 (16.7;60.4)	43.3 (25.0;63.5)	56.8 (33.3;79.2)*	p < 0.01
Work ability score (0–10), mean (SD)	3.5 (2.6)	4.8 (2.5)*	6.8 (2.3)*	3.8 (2.5)	3.1 (2.6)	1.6 (2.0)*	p < 0.01
Receives (or applied for) disability benefits due to LBP, n (%)	22 (25)	35 (17)	7 (5)*	19 (20)	96 (31)*	65 (44)*	p < 0.01

TABLE 3 (*Continued*)

Characteristic	Class 1: Moderate High Disability, Fast Recovery	Class 2: Moderate Disability, Stable	Class 3: Low Moderate Disability, Slow Recovery	Class 4: Moderate High Disability, Late Recovery	Class 5: Moderate High Disability, Stable	Class 6: High Disability, Stable	ANOVA/ Kruskal-Wallis/χ^2
Follow-up							
Treatment(s) at the university hospital during follow-up, n (%)							$p < 0.01$
Multidisciplinary rehabilitation	28 (32)	39 (19)	22 (15)	28 (30)	60 (19)	27 (18)	
Pain anesthesiology treatment	13 (15)	18 (9)	17 (12)	10 (11)	47 (15)	22 (15)	
Surgery	3 (3)	7 (3)	9 (6)	13 (14)*	10 (3)	11 (7)	
Other	0 (0)	0 (0)	0 (0)	1 (1)	2 (1)	0 (0)	
Combination	4 (5)	7 (3)	2 (1)	3 (3)	13 (4)	1 (1)	
No further treatment at the university hospital	39 (45)	135 (66)	98 (66)	38 (41)*	182 (58)	87 (59)	
Very satisfied or satisfied with care, n (%)†	37 (76)*	50 (54)	46 (61)	14 (56)	50 (40)*	21 (40)	$p < 0.01$

EQ5D indicates Eurogol-5D; IQR, interquartile range (Q1-Q3); LBP, low back pain; mo, months; N, number of patients; NIH, National Institutes; ns, not significant; PDI, Pain Disability Index; SD, standard deviation; y, years.

**$P < 0.05$, significantly different mean or median from all other trajectories in post hoc testing (Games-Howell for Anova, Dunn Bonferroni for Kruskall-Wallis), or for ordinal values: significantly different than the expected value for a trajectory (chi-square post hoc test with Bonferroni correction).*

†*58% Missing.*

associated with better or worse treatment outcome by others, but were not associated with trajectories of recovery in the present study.

The number of responders observed during follow-up were illustrative for the different trajectories of disability and LBP impact. At least 75% of patients in recovery groups reached a clinical important outcome at 2 years' follow-up. As mentioned before, recovery-and-relapse on LBP impact was reflected by a sharp decline in responders between 1- and 2-year follow-up. No more than 15% of patients from stable trajectory groups were classified as responder. The small discrepancies seem to be the result of using different methods and thresholds for measuring clinical success. For disability, we utilized an MCIC of 44% improvement compared to baseline, which is higher than the 30% that is often used in LBP studies.[11,39] A lower cut-off would have led to a higher number of responders. It also means that our interpretations reflect conservative estimates.

This study has several limitations. First, data collection took place with patients from a single clinical site, which may challenge generalizability of the results. The impact of LBP on quality of life, work ability, and health care consumption appears higher in this tertiary care sample compared to patients in primary or secondary care.[13] However, pain intensity and disability of GSC patients at baseline are similar to other Dutch LBP and chronic pain samples in primary, secondary, and tertiary care.[13] In this real-world clinical sample, other than having LBP and/or leg pain and an age between 18 and 65 years, there were no strict requirements for participation. Previous research showed that real-world studies with patient populations that are similar to those encountered in clinical practice generally have better external validity than randomized controlled trials.[40] Second, incomplete follow-up was available for most patients. For 29% of patients of the disability sample and 26% of patients of the LBP impact sample, incomplete follow-up was to be expected due to the fact that these patients had not yet reached 24 months of follow-up. Patients with complete follow-up were on average 5 years older than patients with incomplete follow-up. Because the average age was also highest in the more severe and stable trajectories, this could potentially mean that there was an overrepresentation of the number of patients without improvement or recovery. Third, since this study was performed in an uncontrolled longitudinal setting, it is unclear to what extend the identified trajectories were influenced by regression to the mean.[41] Fourth, information on clinical diagnoses was not available. Certain diagnoses and corresponding treatment(s) could be associated with a better or worse prognosis on disability and LBP impact. Data on treatment(s) were limited to those that were provided by the university hospital. Furthermore, the timing of treatment in relation to the assessments varied greatly, and some patients were still receiving (new) treatment after 2 years' follow-up. It would therefore be interesting to re-analyze which trajectories can be identified after 3, 5, or 10 years' follow-up.

TABLE 4. Characteristics of Patients with Different LBP Impact Trajectories.

Characteristic	Class 1: Mild Impact, Recovery	Class 2: Severe Impact Improving to Moderate	Class 3: Severe Impact, Recovery	Class 4: Severe Impact, Stable	Class 5: Moderate Impact, Recovery and Relapse	Class 6: Moderate Impact Improving to Mild	ANOVA/ Kruskal-Wallis/$\chi 2$
N (%)	67 (10)	282 (40)	37 (5)	117 (17)	69 (10)	135 (19)	
Baseline							
Age, y, mean (SD)	44.2 (13.2)	50.3 (12.1)	44.6 (10.4)	51.5 (11.5)	48.7 (12.5)	47.6 (13.0)	$p<0.01$
Sex, n (%) female	33 (49)	165 (59)	17 (46)	71 (61)	40 (58)	74 (55)	ns
Current smoker, n (%)	8 (12)	90 (32)	7 (19)	48 (41)*	14 (20)	19 (14)*	$p<0.01$
Education level, n (%)							$p<0.01$
Low	10 (15)*	118 (42)*	12 (32)	47 (38)	19 (28)	39 (29)	
Middle	22 (33)	81 (29)	13 (35)	46 (39)	28 (41)	50 (37)	
High	31 (46)*	60 (21)	9 (24)	16 (14)	16 (23)	38 (28)	
Other/unknown	4 (6)	23 (8)	3 (8)	8 (7)	6 (9)	8 (6)	
Work Status, n (%)							$p<0.01$
Working	41 (61)*	56 (20)*	11 (30)	11 (9)*	23 (33)	49 (36)	
Partial sick leave	12 (18)	51 (18)	5 (14)	11 (9)	15 (22)	28 (21)	
Sick leave	2 (4)*	48 (17)	14 (38)*	25 (20)	11 (16)	15 (11)	
Unemployed	12 (18)*	127 (45)	7 (19)	70 (60)*	20 (29)	43 (32)	
Duration LBP, n (%)							$p=0.02$
<3 mo	1 (2)	8 (3)	1 (3)	2 (2)	0 (0)	3 (2)	
3 mo–1 y	15 (22)	35 (13)	8 (22)	13 (11)	10 (14)	13 (10)	
1 y–5 y	32 (48)	84 (30)	14 (38)	38 (33)	29 (42)	56 (42)	
>5 y	19 (28)*	155 (55)	14 (38)	64 (55)	30 (44)	63 (47)	
Previous treatment(s) for LBP, n (%)							
Low back surgery	5 (8)*	80 (28)	10 (27)	56 (48)*	14 (20)	28 (21)	$p<0.01$
Opioids	15 (22)*	167 (59)	21 (57)	92 (79)*	28 (41)	61 (45)	$p<0.01$
Injections	4 (6)*	68 (24)	6 (16)	48 (41)*	13 (19)	28 (21)	$p<0.01$
Exercise therapy	56 (84)	249 (88)	34 (92)	105 (90)	62 (90)	122 (90)	ns
Psychological counseling	2 (3)*	55 (20)	1 (3)	31 (27)*	5 (7)	15 (11)	$p<0.01$
Pain intensity (0–10), mean (SD)	5.3 (1.5)*	7.0 (1.5)	6.8 (1.6)	7.7 (1.5)*	6.8 (1.3)	6.0 (1.5)	$p<0.01$
PDI (0–70), mean (SD)	19.2 (11.0)*	40.0 (10.4)	40.7 (12.8)	49.4 (9.9)*	38.1 (12.7)	30.7 (11.4)*	$p<0.01$
Impact Stratification (8–50), mean (SD)	23.1 (4.6)*	36.8 (4.8)	37.4 (4.8)	42.6 (4.2)*	34.8 (5.7)	30.5 (5.8)*	$p<0.01$
EQ5D utility score (−0.33 to 1.00), median (IQR)	0.78 (0.69;0.81)*	0.30 (0.1;0.69)	0.65 (0.30;0.69)	0.19 (0.09;0.30)*	0.60 (0.22;0.75)	0.69 (0.33;0.78)	$p<0.01$
NIH minimal dataset: depression and catastrophizing (0–100), median (IQR)	8.3 (0;25.0)	41.7 (20.8;62.5)	33.3 (20.8;43.8)	58.3 (33.3;79.2)*	33.3 (16.7;52.1)	20.8 (4.2;33.3)	$p<0.01$

TABLE 4 (*Continued*)

Characteristic	Class 1: Mild Impact, Recovery	Class 2: Severe Impact Improving to Moderate	Class 3: Severe Impact, Recovery	Class 4: Severe Impact, Stable	Class 5: Moderate Impact, Recovery and Relapse	Class 6: Moderate Impact Improving to Mild	ANOVA/ Kruskal-Wallis/$\chi 2$
Work ability score (0–10), mean (SD)	6.9 (2.3)*	3.4 (2.6)	4.1 (3.0)	1.6 (2.1)*	4.4 (2.5)	4.8 (2.7)	$p < 0.01$
Receives (or applied for) disability benefits due to LBP, n (%)	3 (5)*	75 (27)	7 (19)	48 (41)*	12 (17)	24 (18)	$p < 0.01$
Follow-up							
Treatment(s) at the university hospital during follow-up, n (%)							$p < 0.01$
Multidisciplinary rehabilitation	11 (16)	60 (21)	11 (30)	18 (15)	26 (38)*	33 (24)	
Pain anesthesiology treatment	7 (10)	42 (15)	3 (8)	24 (21)	7 (10)	16 (12)	
Surgery	5 (8)	13 (5)	6 (16)	6 (5)	6 (9)	5 (4)	
Other	0 (2)	2 (1)	0 (0)	0 (0)	1 (1)	0 (0)	
Combination	0 (0)	12 (4)	3 (8)	5 (4)	0 (0)	4 (3)	
No further treatment at the university hospital	44 (66)	153 (54)	14 (38)	64 (55)	29 (42)	77 (57)	
Very satisfied or satisfied with care, n (%)†	23 (74)	53 (49)	8 (57)	20 (48)	17 (61)	42 (61)	ns

EQ5D indicates Euroqol-5D; IQR, interquartile range (Q1-Q3); LBP, low back pain; mo, months; N, number of patients; NIH, National Institutes; ns, not significant; PDI, Pain Disability Index; SD, standard deviation; y, years.

**$P < 0.05$, significantly different mean or median from all other trajectories in post hoc testing (Games-Howell for Anova, Dunn Bonferroni for Kruskall-Wallis), or for ordinal values: significantly different than the expected value for a trajectory (chi-square post hoc test with Bonferroni correction).*

†59% missing.

TABLE 5. Relationship Between Different Trajectories of LBP Impact and Disability and Response at Follow-up

	% Responders			
	6 mo	12 mo	18 mo	24 mo
Disability				
Total sample	21	31	30	32
Class 1: Moderate high disability, fast recovery	88	97	98	88
Class 2: Moderate disability, stable	19	14	11	15
Class 3: Low moderate disability, slow recovery	47	72	64	76
Class 4: Moderate high disability, late recovery	4	74	85	78
Class 5: Moderate high disability, stable	3	3	1	9
Class 6: High disability, stable	2	0	0	0
LBP impact				
Total sample	n/a	29	n/a	28
Class 1: Mild impact, recovery	n/a	68	n/a	75
Class 2: Severe impact improving to moderate	n/a	7	n/a	11
Class 3: Severe impact, recovery	n/a	100	n/a	100
Class 4: Severe impact, stable	n/a	1	n/a	2
Class 5: Moderate impact, recovery and relapse	n/a	87	n/a	22
Class 6: Moderate impact improving to mild	n/a	22	n/a	42

LBP indicates low back pain; mo, months; n/a, not assessed.

Knowledge on trajectories of functional outcome could potentially help identify clinically relevant subgroups of patients that are (less) likely to respond to LBP treatment. Most patients in this study belonged to subgroups that experienced little to no improvement during follow-up. This shows that there is need for better treatments and/or treatment selection (what works for whom?) for patients with chronic LBP. Alternatively, patients that present with stable trajectories of pain and disability could benefit from improved pain management and coping strategies and patients at risk of recovery and relapse could benefit from effective strategies to prevent recurrence. However, we could not clearly distinguish patients belonging to subgroups with successful outcome from patients with unsuccessful outcome using the characteristics measured in this study. Future research is required to identify which patients are likely to recover or not and whether targeting treatments based on trajectory membership leads to improved outcome.

CONCLUSION

On average, patients show moderate improvements in disability and LBP impact 2 years after visiting a multidisciplinary tertiary spine center. However, latent class analyses revealed that less than one-third of patients belong to recovery trajectory groups and that most patients belong to subgroups experiencing stable levels of disability and LBP impact. Patient characteristics were mostly associated with baseline levels of functioning and did not distinguish recovery trajectories from stable trajectories. Finally, the number of patients meeting thresholds for clinically important change during follow-up align with the different trajectories of disability and LBP impact.

➢ Key Points

- ❑ We identified treatment response trajectories in 996 patients with LBP during and after multidisciplinary care in a tertiary spine center.
- ❑ Using latent class analyses, six disability and six LBP impact trajectories were identified. Most patients belonged to subgroups experiencing stable levels of improvement.
- ❑ Differences in baseline patient characteristics were mostly associated with baseline levels of functioning, instead of (un)favorable outcome during follow-up.
- ❑ Both the average course of functioning for a group of patients, and a single time-point measure of success for individual patients, can present an inaccurate picture of treatment response.

Supplemental digital content is available for this article. Direct URL citations appearing in the printed text are provided in the HTML and PDF version of this article on the journal's Web site (www.spinejournal.com).

References

1. Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. *BMJ* 2006;332:1430–4.
2. Kamper SJ, Apeldoorn AT, Chiarotto A, et al. Multidisciplinary biopsychosocial rehabilitation for chronic low back pain: Cochrane systematic review meta-analysis. *BMJ* 2015;350:h444.
3. O'Keeffe M, Purtill H, Kennedy N, et al. Comparative effectiveness of conservative interventions for nonspecific chronic spinal pain: Physical, behavioral/psychologically informed, or combined? A systematic review and meta-analysis. *J Pain* 2016;17:755–74.

4. Hartvigsen J, Hancock MJ, Kongsted A, et al. What low back pain is and why we need to pay attention. *Lancet* 2018;391:2356–67.
5. Andrew Moore R. What works for whom? Determining the efficacy and harm of treatments for pain. *Pain* 2013;154 (suppl 1):S77–86.
6. Henschke N, van Enst A, Froud R, et al. Responder analyses in randomised controlled trials for chronic low back pain: an overview of currently used methods. *Eur Spine J* 2014;23:772–8.
7. Kongsted A, Kent P, Axen I, et al. What have we learned from ten years of trajectory research in low back pain?. *BMC Musculoskelet Disord* 2016;17:220.
8. Enthoven WT, Koes BW, Bierma-Zeinstra SM, et al. Defining trajectories in older adults with back pain presenting in general practice. *Age Ageing* 2016;45:878–83.
9. Chen Y, Campbell P, Strauss VY, et al. Trajectories and predictors of the long-term course of low back pain: cohort study with 5-year follow-up. *Pain* 2018;159:252–60.
10. O'Brien EM, Staud RM, Hassinger AD, et al. Patient-centered perspective on treatment outcomes in chronic pain. *Pain Med* 2010;11:6–15.
11. Deyo RA, Bryan M, Comstock BA, et al. Trajectories of symptoms and function in older adults with low back disorders. *Spine (Phila Pa 1976)* 2015;40:1352–62.
12. Hebert JJ, Abraham E, Wedderkopp N, et al. Patients undergoing surgery for lumbar spinal stenosis experience unique courses of pain and disability: a group-based trajectory analysis. *PLoS One* 2019;14:e0224200.
13. Dutmer AL, Schiphorst Preuper HR, Soer R, et al. Personal and societal impact of low back pain: The Groningen Spine Cohort. *Spine (Phila Pa 1976)* 2019;44:E1443–51.
14. *Internationaal Richtsnoer Voor 'Good Clinical Practice' Voor Het Onderzoek Met Geneesmiddelen; Vertaling Naar De Nederlandse praktijk*. Den Haag: GCP begeleidingscommissie; 2003.
15. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007;370:1453–7.
16. Tait RC, Chibnall JT, Krause S. The pain disability index: psychometric properties. *Elsevier* 1990;40:171–82.
17. Soer R, Reneman MF, Vroomen PC, et al. Responsiveness and minimal clinically important change of the pain disability index in patients with chronic back pain. *Spine (Phila Pa 1976)* 2012;37:711–5.
18. Soer R, Koke AJ, Vroomen PC, et al. Extensive validation of the pain disability index in 3 groups of patients with musculoskeletal pain. *Spine (Phila Pa 1976)* 2013;38:E562–8.
19. Deyo RA, Dworkin SF, Amtmann D, et al. Report of the NIH task force on research standards for chronic low back pain. *Spine (Phila Pa 1976)* 2014;39:1128–43.
20. Dutmer AL, Reneman MF, Schiphorst Preuper HR, et al. The NIH minimal dataset for chronic low back pain: responsiveness and minimal clinically important change. *Spine (Phila Pa 1976)* 2019;44:E1211–8.
21. Boer A, Dutmer AL, Schiphorst Preuper HR, et al. Measurement properties of the NIH-minimal dataset dutch language version in patients with chronic low back pain. *Spine (Phila Pa 1976)* 2017;42:1472–7.
22. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. the EuroQol group. *Health Policy* 1990;16:199–208.
23. Lamers LM, McDonnell J, Stalmeier PF, et al. The Dutch tariff: Results and arguments for an effective design for national EQ-5D valuation studies. *Health Econ* 2006;15:1121–32.
24. Ilmarinen J, Gould R, Järvikoski A, et al. Diversity of work ability. In: Gould R, Ilmarinen J, Järvisalo J, editors. et al. *Dimensions of Work Ability: Results of the Health 2000 Survey*. Helsinki: Finnish Institute of Occupational Health; 2008; 13-24.
25. Nylund KL, Asparoutiov T, Muthen BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Struct Equ Modeling* 2007; 14:535–69.
26. Geiser C. *Data Analysis with Mplus*. New York: Guilford Publications; 2012.
27. Axén I, Leboeuf-Yde C. Trajectories of low back pain. *Best Pract Res Clin Rheumatol* 2013;27:601–12.
28. da Silva T, Mills K, Brown BT, et al. Recurrence of low back pain is common: a prospective inception cohort study. *J Physiother* 2019; 65:159–65.
29. Adnan R, Van Oosterwijck J, Cagnie B, et al. Determining predictive outcome factors for a multimodal treatment program in low back pain patients: a retrospective cohort study. *J Manipulative Physiol Ther* 2017;40:659–67.
30. Dionne C, Koepsell TD, Von Korff M, et al. Formal education and back-related disability. In search of an explanation. *Spine (Phila Pa 1976)* 1995;20:2721–30.
31. Mallen CD, Peat G, Thomas E, et al. Prognostic factors for musculoskeletal pain in primary care: a systematic review. *Br J Gen Pract* 2007;57:655–61.
32. Hellum C, Johnsen LG, Gjertsen Ø, et al. Predictors of outcome after surgery with disc prosthesis and rehabilitation in patients with chronic low back pain and degenerative disc: 2-year follow-up. *Eur Spine J* 2012;21:681–90.
33. Chou R, Shekelle P. Will this patient develop persistent disabling low back pain?. *JAMA* 2010;303:1295–302.
34. Turner JA, Shortreed SM, Saunders KW, et al. Optimizing prediction of back pain outcomes. *Pain* 2013;154:1391–401.
35. Turner JA, Franklin G, Fulton-Kehoe D, et al. Worker recovery expectations and fear-avoidance predict work disability in a population-based workers' compensation back pain sample. *Spine (Phila Pa 1976)* 2006;31:682–9.
36. Leboeuf-Yde C, Grønstvedt A, Borge JA, et al. The nordic back pain subpopulation program: demographic and clinical predictors for outcome in patients receiving chiropractic treatment for persistent low back pain. *J Manipulative Physiol Ther* 2004;27:493–502.
37. van Hooff ML, Spruit M, O'Dowd JK, et al. Predictive factors for successful clinical outcome 1 year after an intensive combined physical and psychological programme for chronic low back pain. *Eur Spine J* 2014;23:102–12.
38. Edwards RR, Klick B, Buenaver L, et al. Symptoms of distress as prospective predictors of pain-related sciatica treatment outcomes. *Pain* 2007;130:47–55.
39. Ostelo RW, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976)* 2008;33:90–4.
40. Blonde L, Khunti K, Harris SB. Interpretation and impact of real-world clinical data for the practicing clinician. *Adv Ther* 2018;35:1763–74.
41. Ostermann T, Willich SN, Lüdtke R. Regression toward the mean—a detection method for unknown population mean based on Mee and Chua's algorithm. *BMC Med Res Methodol* 2008;8:52.