Patients' treatment beliefs in low back pain: development and validation of a questionnaire in primary care

Published in:
Pain

DOI:
10.1097/j.pain.0000000000000193

Link to publication

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Patients’ treatment beliefs in low back pain: development and validation of a questionnaire in primary care

Alexandra Dima1,1, George T. Lewith2, Paul Little3, Rona Moss-Morris4, Nadine E. Foster5, Matthew Hankins6, George Surtees7, Felicity L. Bishop1,2,*

Abstract
Choosing the most appropriate treatment for individual patients with low back pain (LBP) can be challenging, and clinical guidelines recommend taking into account patients’ preferences. However, no tools exist to assess or compare patients’ views about LBP treatments. We report the development and validation of the Low Back Pain Treatment Beliefs Questionnaire (LBP-TBQ) for use across different treatments in clinical practice and research. Using qualitative data, we developed a pool of items assessing perceived credibility, effectiveness, concerns about, and individual “fit” of specific treatments. These items were included in a survey completed by 429 primary care patients with LBP, of whom 115 completed it again 1 to 2 weeks later. We performed psychometric analyses using nonparametric item response theory and classical test theory. The 4 subscales of the resulting 16-item LBP-TBQ showed good homogeneity (H = 0.46-0.76), internal consistency (α = 0.73-0.94), and stability (r = 0.63-0.83), confirmed most convergent and discriminant validity hypotheses, and had acceptable structural validity for 4 guideline-recommended treatments: pain medication, exercise, manual therapy, and acupuncture. Participants with stronger positive treatment beliefs were more likely to rank that treatment as their first choice, indicating good criterion validity (t values = 3.11-9.80, all P < 0.01, except pain medication effectiveness beliefs, t(339) = 1.35; P = 0.18). A short 4-item version also displayed good homogeneity (H = 0.43-0.66), internal consistency (α = 0.70-0.86), and stability (r = 0.92-0.95) and was significantly related to treatment choice (t values = 4.33-9.25, all P < 0.01). The LBP-TBQ can be used to assess treatment beliefs in primary care patients with LBP and to investigate the effects of treatment beliefs on treatment uptake and adherence.

Keywords: Low back pain, Questionnaire validation, Scale, Psychometrics, Treatment beliefs, Medication beliefs, Pain medication, Exercise, Manual therapy, Acupuncture, Nonparametric item response theory

1. Introduction
Low back pain (LBP) is a leading cause of disability worldwide and is managed mostly within primary care. Most patients have nonspecific LBP, and 75% may continue to have pain and/or disability 12 months after initial consultation. Clinical guidelines recommend several treatments, including conventional (eg, education, exercise, pain medication), complementary/alternative (CAMs) (eg, acupuncture, manual therapy), and combined physical and psychological treatments (pain management courses). The clinical challenge is to choose optimal treatments for individuals; clinical guidelines explicitly encourage considering patients’ preferences but offer no recommendations on how to elicit and integrate them into decision making. Clear conceptualization and standardized assessment of patients’ preferences would facilitate further research and possible subsequent integration into practice.

Treatment preferences can be understood within the extended Common Sense Model (CSM) of illness representations. This model stipulates that when confronted with a medical problem, patients develop cognitive and emotional representations of their condition and beliefs about possible treatments (“treatment beliefs”) based on information from various sources, which guide their behaviours (eg, treatment choice) and can predict subsequent clinical outcomes. Significant relationships have been found between illness representations, treatment beliefs, and outcomes such as adherence and satisfaction in various chronic conditions, including LBP. According to the CSM, treatment preferences develop when patients attempt to “match” treatments to their condition, aiming for coherence between illness representations and treatment beliefs. For example, patients who believe their LBP is caused by a mechanical problem may prefer treatments they believe can...
remedy mechanical dysfunctions and choose manual therapy; patients who see LBP as essentially a pain symptom may prefer treatments they consider appropriate to reduce pain and choose pain medication. Reliable and valid measurement of treatment beliefs in LBP is needed to further test such hypotheses derived from the CSM and facilitate shared decision making.

Illness perceptions have been examined extensively; validated questionnaires are available and have been used in LBP research. However, we could not identify a treatment beliefs questionnaire applicable to different LBP treatments that concomitantly assesses several relevant beliefs. Existing measures are treatment-specific, and previous studies in LBP have focused on single belief dimensions, eg, expectations of effectiveness and perceived credibility. However, qualitative research suggests that LBP treatment beliefs are multidimensional. In our recent qualitative study, patients evaluated LBP treatments according to 4 specific dimensions: perceived credibility, individual fit, concerns, and effectiveness. Here, we report the development and validation of a questionnaire, the Low Back Pain Treatment Beliefs Questionnaire (LBP-TBQ), which assesses patients’ beliefs about 4 practitioner-delivered primary care treatments: pain medication, exercise, manual therapy, and acupuncture. We focused on these treatments, as they are the frontline treatments named in the National Institute for Health and Care Excellence (NICE) care pathway for persistent nonspecific LBP and are also recommended by the American College of Physicians and the American Pain Society LBP guidelines; pain medication, exercise, and manual therapy are also endorsed in the European guidelines for chronic nonspecific LBP.

2. Methods

2.1. Instrument development

In our previous qualitative study, we showed how 75 patients participating in 13 focus groups evaluated specific LBP treatments according to whether they perceived them to be believable and to “make sense” (credibility); expected them to lead to symptom improvements (effectiveness); had concerns that treatments might cause further damage or have side effects (concerns); felt the treatment would be a suitable solution for them personally (individual fit). Because patients expressed these beliefs about specific treatments (eg, pain medication, acupuncture) we termed these Specific Treatment Beliefs. Themes reflecting the context of treatment decision making also emerged and highlighted the importance of understanding patients’ more general treatment-seeking beliefs: their need for a clear diagnosis, their willingness to try different treatments, their interest in self-management, and their expectations regarding the health care system. We developed an item pool comprising 71 items, 27 items assessing the Specific Treatment Beliefs (the focus of this article) and 44 assessing the contextual themes (to be reported elsewhere).

We reviewed our qualitative data to choose item content and wording that reflected topics and terminology used by participants. To facilitate comparisons between patients’ beliefs about different treatments, items assessing Specific Treatment Beliefs were designed to be answered 4 times, once each in relation to pain medication, exercise, manual therapy, and acupuncture. Therefore, items specific to particular treatments (eg, fear of needles in acupuncture) were not included. Remaining items were worded more generally to capture these specific issues (eg, “I have concerns about [acupuncture] for my back pain”).

We pretested the initial pool of 27 items using cognitive “think aloud” interviews with 10 adults with LBP. This pretest allowed us to select the most appropriate items for further testing and to adjust item content and wording to enhance face validity and acceptability. After the pretest, we retained 20 items on Specific Treatment Beliefs in the Draft LBP-TBQ for psychometric testing. The reasons for exclusion were (1) participants interpreted the item in a different way from the intended meaning (1 item), (2) the item was too similar to another item that was perceived as clearer (4 items), (3) the respondents had difficulties applying it to all 4 treatments (1 item), or (4) the item was more related to the context of care than to the treatment itself (1 item). We opted for a lower number of items and a confirmatory approach to psychometric testing (instead of a higher number of items and an exploratory approach) because of the increased patient burden involved in answering questions repeatedly for each of the 4 treatments and to facilitate analysis of the structure of the questionnaire across all 4 treatments.

2.2. Design and procedure

We included the Draft LBP-TBQ, items on the context of treatment decision making, several validating measures, and questions on demographic and clinical characteristics in a self-report survey of adults (at least 18 years) with LBP. We included adults who reported LBP for at least 6 weeks because our previous qualitative work revealed that although the NICE guidelines particularly focus on persistent nonspecific LBP (ie, pain not caused by malignancy, infection, fracture, inflammatory disorders, nerve root compression, and lasting between 6 weeks and 12 months), the distinction between persistent and chronic LBP is rarely used in practice by clinicians or patients. Therefore, we did not apply additional exclusion criteria but compared responses to our questionnaire across different subgroups of patients.

Participants were recruited between November 2011 and March 2012 from public sector primary care physicians (general practitioners) and private sector CAM clinics in 3 South England counties (Hampshire, Wiltshire, and Dorset), and advertisements on online U.K.-based patient forums. We aimed for 400 participants, a statistically acceptable sample size for our planned psychometric analyses, acknowledging that statistical power also depends on data properties that could not be estimated before analysis. Physicians and CAM clinicians forwarded paper-based surveys to their eligible patients by post. Online advertisements linked directly to an identical web-based survey. To enable examination of test–retest reliability, participants were asked to volunteer to complete the LBP-TBQ again; all such volunteers were sent a second survey by post or e-mail 1 week later. We obtained ethics approval from Southampton and South West Hampshire REC B (10/H0504/78).

2.3. Draft Low Back Pain Treatment Beliefs Questionnaire

In the 20-item Draft LBP-TBQ, 4 items assessed perceptions of credibility (2 negatively worded, ie, described in terms of doubting the credibility of the treatment), 5 items assessed perceived effectiveness (2 negatively worded), 8 items assessed concerns (4 negatively worded) and 5 items assessed perceived individual fit (3 negatively worded) (see Table 2 for item content). A 5-point verbal response scale was used for all items (Strongly Disagree,
Disagree, Neither Agree nor Disagree, Agree, Strongly Agree), and scored (1-5) such that a high score represented positive beliefs about the LBP treatment. Each set of 20 items was presented with respect to each of the 4 treatments: pain medication, exercise, manual therapy, and acupuncture (so each participant responded to 80 LBP-TBQ items). Definitions of these treatments, based on the U.K. NICE guidelines, were provided to limit variability in interpreting treatment labels and encourage answers that can be interpreted within the context of U.K. clinical practice; these definitions may need to be adjusted for different purposes in future research (subject to confirmation of psychometric properties in specific other contexts and populations).

2.4. Validating measures

We developed hypotheses about relationships between each validating measure and the Draft LBP-TBQ subscales (see below and Table 3). In brief, to demonstrate convergent validity, we required at least medium or strong significant correlations (ie, \( r \geq 0.3 \)). To demonstrate discriminant validity, we required at most small-to-moderate significant correlations (\( r < 0.3 \)).

2.4.1. Beliefs About Medicines Questionnaire—general beliefs

Respondents completed 5-point scales to rate their agreement with statements representing beliefs about the potential harmful effects of medicines (Beliefs About Medicines Questionnaire [BMQ] Harm, 4 items, eg, "medicines do more harm than good," Cronbach \( \alpha = 0.68 \) in the present sample) and about medicines being overprescribed by doctors (BMQ Overuse, 4 items, eg, "doctors use too many medicines," \( \alpha = 0.76 \)). High scores indicated more negative beliefs about medicines. Both BMQ scales were used to assess the convergent validity of the LBP-TBQ Concerns subscale for pain medication.

2.4.2. The Brief Illness Perception Questionnaire

Single items with 11-point response scales assessed 8 dimensions of illness perceptions: consequences (the extent to which LBP affects one’s life), timeline (the expected duration of LBP), personal control (the extent to which one perceives control over one’s LBP), treatment control (the extent to which one perceives one’s treatment controls one’s LBP), identity (the number of symptoms associated with LBP), coherence (the extent to which one understands one’s LBP), concern (the extent of concerns about LBP), and emotional response (the extent to which movement leads to (re)injury and should therefore be avoided). The Tampa Scale of Kinesiophobia Activity Avoidance subscale was used to assess the convergent validity of the LBP-TBQ Concerns subscale for pain medication.

2.4.3. Credibility Expectancy Questionnaire

Two subscales assessed perceptions of treatment credibility (Credibility Expectancy Questionnaire [CEQ] Credibility, 3 items with 9-point response scales, eg, "At this point, how logical does [treatment] seem?") and outcome expectancy (CEQ Expectancy, 1 item with 9-point response scales and 2 items with a 11-point response scale, from 0% to 100%, eg, "By the end of a course of [treatment], how much improvement in your back pain do you think would occur?"). To reduce response burden, each respondent answered the CEQ in relation to 1 of the 4 treatments only (randomised allocation). Good internal consistency was shown in our sample for credibility (\( \alpha \) range: 0.85-0.94) and expectancy (\( \alpha \) range: 0.85-0.96) scales for all treatments. High scores indicated perceiving the treatment as more believable, convincing, and logical and as leading to bigger improvements. The CEQ Credibility was used to assess the convergent validity of the LBP-TBQ credibility subscale for all 4 treatments. The CEQ Expectancy was used to assess the convergent validity of the LBP-TBQ expectancy subscale for all 4 treatments.

2.4.4. Holistic Complementary and Alternative Medicine Questionnaire Attitudes to CAM Subscale

Six statements assessed general attitudes towards CAM using 6-point agree/disagree response scales (eg, "It is worthwhile trying complementary medicine before going to the doctor," \( \alpha = 0.71 \)). High scores indicated stronger beliefs that CAM is ineffective and unscientific compared with mainstream medicine. The Holistic Complementary and Alternative Medicine Questionnaire Attitudes to CAM subscale was used to assess the convergent validity of the LBP-TBQ subscales for all 4 treatments.

2.4.5. Tampa Scale of Kinesiophobia (TSK-11) Activity Avoidance Subscale

Six items assessed beliefs about the relationship between movement, pain, and reinjury (Tampa Scale of Kinesiophobia Activity Avoidance, eg, "I’m afraid I might injure myself if I exercise," \( \alpha = 0.81 \)) using 4-point agree-disagree response scales. High scores indicated more intense concerns and beliefs that movement leads to (re)injury and should therefore be avoided. The Tampa Scale of Kinesiophobia Activity Avoidance subscale was used to assess the convergent validity of the LBP-TBQ concerns subscale for exercise.

2.4.6. Treatment ranking

To assess criterion validity, participants were asked to rank the 4 treatments (pain medication, exercise, manual therapy, acupuncture) in order of preference, starting with the treatment they would most like to have. Choices regarding each treatment were coded separately and dichotomised to identify 2 groups of patients for each treatment: those who did and those who did not select that treatment as their first choice. We hypothesised that scores on the LBP-TBQ subscales would distinguish between these 2 groups of patients for each treatment.
2.5. Participant demographic and clinical characteristics

We used single items to assess sociodemographic characteristics (age, gender, education level, employment status, ethnicity) and clinical characteristics (duration of LBP, previous/current use of various treatments for LBP and satisfaction with these treatments, self-reported comorbidities including sciatica, receiving benefits for LBP, perception of general health status).

Low back pain severity over the past 6 months was assessed with the Chronic Pain Grade Questionnaire, which comprises 3 pain intensity items (present, worst, and average) and 4 disability items (number of days kept from usual activities and pain interference with daily activities, recreational or social activities, and work). Responses were used to compute 5 pain grades, from no pain and disability to the highest disability irrespective of pain intensity. Participants who had experienced LBP in the past year (single item) were asked whether they currently experience pain and whether they had experienced 3 symptoms suggestive of a differential diagnosis of nerve root compression: leg pain worse than back pain, leg pain worse when coughing or sneezing, and numbness or pins and needles in the leg or feet. Participants who reported any of these symptoms were classified as more likely to have nerve root compression.

2.6. Data analysis

Data analysis was performed in SPSS Statistics 21 (IBM SPSS Statistics for Windows, version 21.0, released 2012; IBM Corp, Armonk, NY), Amos 21, and R (Mokken package). Data entry was checked for accuracy. We identified no systematic pattern of missing data for the selected variables (Little Missing Completely At Random, MCAR test; \( \chi^2(36,655) = 36,771.69, P = 0.33 \)). Missing data were computed using expectation maximization for the relevant questionnaire items (excluding items where noncompletion was expected based on responses to previous items). We computed descriptive statistics for demographics, clinical data, and validation measures. We performed psychometric analyses separately for specific and general treatment beliefs items.

2.6.1. Item selection and structural validity

We used a confirmatory approach to questionnaire validation and aimed for a final item set that included an optimal, parsimonious, number of items that would permit the use of both sum scores and individual items in latent variable models (ie, 4 items per construct, of which 2 should be negatively worded). We examined the structure of the item sets on the first-wave data using nonparametric item response theory, ie, Mokken scale analysis (MSA). According to MSA, items order respondents stochastically on one latent dimension representing the target construct if they meet 3 criteria: (1) unidimensionality (ie, respondents that endorse more intense items are also more likely to endorse less intense items, while endorsing less intense items is not related to the probability of endorsing more intense items), (2) local independence (ie, the statistical relationship between items should be explained solely by the latent construct), and (3) latent monotonicity (ie, the probability of endorsing an item should not decrease with increasing levels of the latent construct).

We investigated these properties by calculating coefficients of homogeneity (H) at item, item-pair, subscale, and scale level. Homogeneity values range from 0 (no association) to 1 (perfect association given differences in item intensity), where 0.3 to 0.4 indicates weak, 0.4 to 0.5 medium, and values above 0.5 good homogeneity. To reduce scale length, items presenting low homogeneity and violations of latent monotonicity were considered for exclusion. We subsequently examined correlations between the resulting subscales.

We further investigated structural validity using confirmatory factor analyses (CFAs), as MSA is nonparametric and therefore does not allow modelling the effect of using positively and negatively worded items. Models were evaluated in relation to established criteria for the likelihood ratio \( \chi^2 \) test, incremental fit, and residual-based indices (TLI and CFI > 0.95; RMSEA and \( \chi^2 \); \( P \) value < 0.05). We followed recommendations to judge model fit flexibly within the broader context of model diagnosis and theory and to consider fit criteria alongside overall model tenability and possible sources of misspecification.

To investigate alternative structures, we used an automated item selection procedure to group items into Mokken scales in a data-driven manner and identify unscaleable items at increasing homogeneity threshold values.

2.6.2. Reliability

The new subscales were also examined according to classical test theory. Internal consistency was assessed using Cronbach \( \alpha \) (above 0.70), item-total correlations, and Cronbach \( \alpha \) if item excluded. Correlations between responses to the same scales at the first and second survey administration were used to judge test–retest reliability over 1 to 2 weeks based on a threshold of \( r = 0.70 \).

2.6.3. Convergent and discriminant validity

The subscales were tested for convergent and discriminant validity against existing measures using Pearson correlations (sensitivity analyses were performed with Spearman \( \rho \) and Kendall \( \tau \), to account for the ordinal level of some measures).

2.6.4. Criterion validity

The new LBP-TBQ subscales were examined against treatment ranking reports to assess their ability to predict treatment preferences. We used \( t \) tests to compare scores between participants who did and did not rank each treatment as their first choice.

2.6.5. Measurement invariance

Multigroup CFAs were performed to investigate whether different subgroups of respondents attribute the same meaning to the target construct (metric invariance) and whether respondents with equal scores on the latent construct also have equal scores on the items (scalar invariance). If a scale has these properties, then group differences in mean scores can be interpreted as substantive differences, as they are not due to participants in different groups attributing different meanings to the scale or to measurement bias of individual items in these groups. Five comparisons were considered: respondents with nerve root compression likely or unlikely (based on self-reported presence of 1 or more of 3 indicative symptoms), respondents who self-reported sciatica diagnosed in relation to LBP or not, respondents with or without experience of each treatment, and the 2 data collection waves.

2.6.6. Selection and validation of short-form Low Back Pain Treatment Beliefs Questionnaire

To increase the feasibility of using the LBP-TBQ in multimeasure patient surveys and fast-paced clinical contexts, we developed a short 4-item LBP-TBQ version by selecting 1 best-performing...
3. Results

3.1. Participants

The survey was completed by 429 participants, of whom 344 (80%) responded to the 1498 invitations mailed to public sector physicians’ patients (23% response rate). Participants were aged 18 to 90 years (mean = 55; SD = 15.2); 247 (60%) were women, and 393 (91.6%) were of British, Irish, or other white ethnic background. The majority (335; 78.1%) completed the paper version. Participants reported having LBP for between 6 weeks and 52 years (median: 6 years; interquartile range: 13.18 years); 415 (96%) had LBP in the past year, 400 (93.2%) in the past 6 months, and 308 (71.8%) at the time of survey; 398 (92.8%) considered their general health to be “fair” to “very good,” but more than half reported high disability (ie, chronic pain grade III or IV). Only 81 (15.5%) reported receiving state benefits for LBP. Of the 174 volunteering to complete the survey twice, 115 (66.1%) participants completed and returned the LBP-TBQ again 1 to 2 weeks later. There were no differences between respondents and nonrespondents to the second survey in age, gender, pain duration, general health levels, pain intensity, disability levels, or chronic pain grade. Additional sociodemographic and clinical data and descriptive statistics for the validating measures are available in Table 1 and in Supplemental Digital Content 1 (available online at http://links.lww.com/PAIN/A79).

3.2. Item selection and structural validity

Items were selected iteratively based on homogeneity and monotonicity results at item, item-pair, and scale level for all 4 treatments. Four items were excluded based on item properties and content to achieve four 4-item subscales with 2 reversed items each and good homogeneity and monotonicity, except 3 significant violations of monotonicity for acupuncture items (Table 2; see also Supplemental Digital Content 2, which shows initial homogeneity values for all items and violations of monotonicity for acupuncture items [available online at http://links.lww.com/PAIN/A80]).

We examined structural validity through CFA for each treatment by comparing the hypothesized 4-factor model (covariance between factors and covariances between error terms of reverse-coded items) with several alternatives (1-factor model, 4-factor model with 1 common higher order factor, 1-factor model improved by specifying error covariances suggested by modification indices). Although none of the models reached threshold values for all fit indices and all treatment types, the 4-factor model performed slightly better than its alternatives. As an example, Figure 1 displays model parameter estimates for pain medication items (Supplemental Digital Content 3 shows parameter estimates and model fit for alternative models and other treatment types [available online at http://links.lww.com/PAIN/A81]).

3.2.1. Reliability

All 4 subscales showed acceptable internal consistency, with Cronbach α values ranging from 0.73 to 0.94. They also showed acceptable test–retest validity; over 1 to 2 weeks, pain medication effectiveness exhibited the lowest stability (r = 0.63) and manual therapy concerns exhibited the highest stability (r = 0.83). For details, see Table 3.

3.3. Convergent and discriminant validity

As hypothesized, on the whole, the LBP-TBQ subscales were at least moderately associated with conceptually related constructs (r > 0.3) and showed medium to nonsignificant associations (r < 0.3) with constructs expected to be conceptually different (Table 4). The main exceptions involved the discriminant validity of the exercise subscales, which unexpectedly correlated with LBP perceptions (Brief Illness Perception Questionnaire); exercise concerns and exercise individual fit were moderately correlated with LBP concerns; and exercise concerns were moderately correlated with perceived emotional impact of LBP. Related to convergent validity, the observed correlations of 0.29 fell just short of the hypothesized 0.3 between attitudes to CAM (Holistic Complementary and Alternative Medicine Questionnaire) and beliefs regarding effectiveness and individual fit of manual therapy and between general beliefs about the harmful effects of medications (BMQ) and LBP-specific medication concerns.

### Table 1

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N (%)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>Secondary school or less</td>
<td>183 (42.7)</td>
</tr>
<tr>
<td>Sixth form college</td>
<td>107 (24.9)</td>
</tr>
<tr>
<td>Undergraduate study</td>
<td>75 (17.5)</td>
</tr>
<tr>
<td>Postgraduate study</td>
<td>34 (7.9)</td>
</tr>
<tr>
<td>Work status</td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>179 (41.7)</td>
</tr>
<tr>
<td>At usual job</td>
<td>152 (35.4)</td>
</tr>
<tr>
<td>On light duty</td>
<td>16 (3.7)</td>
</tr>
<tr>
<td>Paid leave or sick leave</td>
<td>8 (1.9)</td>
</tr>
<tr>
<td>Unpaid leave</td>
<td>3 (0.7)</td>
</tr>
<tr>
<td>Retired</td>
<td>134 (31.2)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>230 (53.1)</td>
</tr>
<tr>
<td>Because of LBP</td>
<td>25 (5.8)</td>
</tr>
<tr>
<td>On disability</td>
<td>22 (5.1)</td>
</tr>
<tr>
<td>Homemaker</td>
<td>27 (6.3)</td>
</tr>
<tr>
<td>Student</td>
<td>9 (2.1)</td>
</tr>
<tr>
<td>For other reasons</td>
<td>13 (3.0)</td>
</tr>
<tr>
<td>Pain duration (3 categories)</td>
<td></td>
</tr>
<tr>
<td>Persistent LBP (6 wk-12 mo)</td>
<td>88 (20.5)</td>
</tr>
<tr>
<td>Chronic/recurrent LBP (12 mo-3 y)</td>
<td>63 (14.7)</td>
</tr>
<tr>
<td>Chronic/recurrent LBP (&gt;3 yrs)</td>
<td>278 (64.8)</td>
</tr>
<tr>
<td>Chronic pain grade</td>
<td></td>
</tr>
<tr>
<td>Grade I: low disability, low intensity</td>
<td>82 (19.1)</td>
</tr>
<tr>
<td>Grade II: low disability, high intensity</td>
<td>90 (21)</td>
</tr>
<tr>
<td>Grade III: high disability, moderately limiting</td>
<td>81 (18.9)</td>
</tr>
<tr>
<td>Grade IV: high disability, severely limiting</td>
<td>147 (34.3)</td>
</tr>
<tr>
<td>Reporting one or more comorbidities</td>
<td>282 (65.7)</td>
</tr>
<tr>
<td>Subgroups examined for MI</td>
<td></td>
</tr>
<tr>
<td>Pain duration &lt;3 y</td>
<td>151 (35.2)</td>
</tr>
<tr>
<td>Self-reported sciatica</td>
<td>192 (44.8)</td>
</tr>
<tr>
<td>At least 1 symptom of nerve root compression</td>
<td>144 (33.6)</td>
</tr>
<tr>
<td>Past experience of medication</td>
<td>208 (48.5)</td>
</tr>
<tr>
<td>Past experience of manual therapy</td>
<td>273 (63.6)</td>
</tr>
<tr>
<td>Past experience of exercise</td>
<td>187 (43.6)</td>
</tr>
<tr>
<td>Past experience of acupuncture</td>
<td>127 (29.6)</td>
</tr>
</tbody>
</table>

* Percentages reported without including missing values.
LBP, low back pain; MI, measurement invariance.
3.4. Criterion validity

For each treatment, those participants who ranked a treatment as their top choice had more positive beliefs about that treatment than did other participants. This difference was significant for all treatments and all subscales, except beliefs about the effectiveness of pain medication (Table 5). In other words, when participants had stronger beliefs about a treatment’s effectiveness, credibility, and individual fit and had fewer concerns about a treatment, they were more likely to prefer that treatment.

3.5. An alternative structure

Subscale scores were highly correlated with each other within treatments (shown in Supplemental Digital Content 4, available online at http://links.lww.com/PAIN/A82). An exploratory MSA (automated item selection procedure) (results shown in Supplemental Digital Content 5, available online at http://links.lww.com/PAIN/A83) revealed that the entire item set could alternatively be considered a single 16-item scale with medium to good homogeneity. For this global scale,
homogeneity scores were 0.46, 0.59, 0.60, 0.67 and Cronbach α values were 0.92, 0.95, 0.95, and 0.96 for pain medication, acupuncture, exercise, and manual therapy, respectively.

3.6. Measurement invariance

Multigroup CFAs (summarised in Table 6) indicated that all scales showed scalar invariance between the 2 data collection waves and in most other subgroup comparisons, with some exceptions. Manual Therapy Beliefs displayed metric invariance between subgroups differentiated on probability of nerve root compression and self-reported sciatica diagnoses (results shown in Supplemental Digital Content 6, available online at http://links.lww.com/PAIN/A84). Exercise Beliefs displayed metric invariance between treatment experience subgroups (results shown in Supplemental Digital Content 7, available online at http://links.lww.com/PAIN/A85). No measurement invariance (MI) was observed for Medication Beliefs between patients with and without nerve compression (results shown in Supplemental Digital Content 8, available online at http://links.lww.com/PAIN/A86). No MI was observed for Acupuncture Beliefs in patients with or without sciatica and with or without treatment experience (results shown in Supplemental Digital Content 9, available online at http://links.lww.com/PAIN/A87). In addition, no MI was found between medication, exercise, manual therapy, and acupuncture regarding the LBP-TBQ scales (results shown in Supplemental Digital Content 10, available online at http://links.lww.com/PAIN/A88).
Table 4
Convergent and discriminant validity for each treatment version of LBP-TBQ subscales.

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Pain medication</th>
<th>Exercise</th>
<th>Manual therapy</th>
<th>Acupuncture</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hypothesis</td>
<td>Pearson r</td>
<td>Hypothesis</td>
<td>Pearson r</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Credibility</td>
<td>↑ CEQ Credibility 0.44*</td>
<td>↑ CEQ Credibility 0.66*</td>
<td>↑ CEQ Credibility 0.67*</td>
<td>↑ CEQ Credibility 0.68*</td>
</tr>
<tr>
<td></td>
<td>= BIPQ Concerns 0.05</td>
<td>= BIPQ Concerns 0.29*</td>
<td>= BIPQ Concerns 0.22*</td>
<td>= BIPQ Concerns 0.02</td>
</tr>
<tr>
<td></td>
<td>= BIPQ Emotions 0.06</td>
<td>= BIPQ Emotions 0.24*</td>
<td>= BIPQ Emotions 0.18*</td>
<td>= BIPQ Emotions 0.02</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>↑ CEQ Expectancy 0.48*</td>
<td>↑ CEQ Expectancy 0.59*</td>
<td>↑ CEQ Expectancy 0.53*</td>
<td>↑ CEQ Expectancy 0.50*</td>
</tr>
<tr>
<td></td>
<td>= BIPQ Concerns 0.15*</td>
<td>= BIPQ Concerns 0.27*</td>
<td>= BIPQ Concerns 0.24*</td>
<td>= BIPQ Concerns 0.03</td>
</tr>
<tr>
<td></td>
<td>= BIPQ Emotions 0.13*</td>
<td>= BIPQ Emotions 0.22*</td>
<td>= BIPQ Emotions 0.19*</td>
<td>= BIPQ Emotions 0.05</td>
</tr>
<tr>
<td>(Fewer) Concerns</td>
<td>↓ BMQ Harm 0.29*</td>
<td>↓ TSK Activity Avoidance 0.60*</td>
<td>↓ HCAMQ Attitudes 0.32*</td>
<td>↓ HCAMQ Attitudes 0.36*</td>
</tr>
<tr>
<td></td>
<td>↓ BMQ Overuse 0.39*</td>
<td>= BIPQ Concerns 0.39*</td>
<td>= BIPQ Concerns 0.27*</td>
<td>= BIPQ Concerns 0.05</td>
</tr>
<tr>
<td></td>
<td>= BIPQ Concerns 0.10*</td>
<td>= BIPQ Emotions 0.38*</td>
<td>= BIPQ Emotions 0.27*</td>
<td>= BIPQ Emotions 0.10†</td>
</tr>
<tr>
<td></td>
<td>= BIPQ Emotions 0.11†</td>
<td>= BIPQ Emotions 0.11†</td>
<td>= BIPQ Emotions 0.11†</td>
<td>= BIPQ Emotions 0.11†</td>
</tr>
<tr>
<td>Individual fit</td>
<td>= BIPQ Concerns 0.04</td>
<td>= BIPQ Concerns 0.32*</td>
<td>↓ HCAMQ Attitudes 0.29*</td>
<td>↓ HCAMQ Attitudes 0.31*</td>
</tr>
<tr>
<td></td>
<td>= BIPQ Emotions 0.08</td>
<td>= BIPQ Emotions 0.27**</td>
<td>= BIPQ Concerns 0.23*</td>
<td>= BIPQ Concerns 0.04</td>
</tr>
<tr>
<td></td>
<td>= BIPQ Emotions 0.11†</td>
<td>= BIPQ Emotions 0.11†</td>
<td>= BIPQ Emotions 0.23*</td>
<td>= BIPQ Emotions 0.05</td>
</tr>
</tbody>
</table>

† indicates at least moderate positive correlation hypothesised (convergent validity; r ≥ 0.3); ↓ indicates at least moderate negative correlation hypothesised (convergent validity; r ≥ 0.3); † indicates at most moderate correlation hypothesised (divergent validity; r < 0.3).

* P < 0.01.
† P < 0.05.
BMQ, Beliefs About Medicines Questionnaire; CEQ, Credibility Expectancy Questionnaire; HCAMQ, Holistic Complementary and Alternative Medicine Questionnaire; LBP-TBQ, Low Back Pain Treatment Beliefs Questionnaire; TSK, Tampa Scale of Kinesiophobia.

3.7. Short version of Low Back Pain Treatment Beliefs Questionnaire

One item was selected to represent each subscale based on scores on homogeneity and Cronbach α if item excluded (the item with best properties for all treatments). The 4-item version of the questionnaire displayed good homogeneity (H = 0.43-0.66) (Table 7), internal consistency (α = 0.70-0.86), and stability (r = 0.82-0.85) (Table 8). People who ranked a treatment as their first choice reported significantly more positive treatment beliefs than people who ranked the treatment as a second, third, or last option (Table 9).

4. Discussion

To facilitate the integration of treatment preferences in LBP clinical decision making in primary care and to stimulate further research on this topic, we developed and validated a 16-item scale to measure core beliefs about 4 recommended LBP treatments (pain medication, exercise, manual therapy, acupuncture). The newly developed LBP-TBQ showed good item properties, homogeneity, internal consistency, and stability. Discriminant and convergent validity were confirmed for most treatments, and the 4-factor structure was largely confirmed. Participants were more likely to rank a treatment as their first...
choice if they had stronger beliefs about its effectiveness, credibility, and individual fit, and fewer concerns about it, thus supporting the criterion validity of LBP-TBQ. A short 4-item LBP-TBQ was also developed with best-performing items and showed good psychometric properties. Both 16-item and 4-item versions (shown in Supplemental Digital Content 11, available online at http://links.lww.com/PAIN/A89) can be used in future research and clinical practice to assess patients’ beliefs about treatments.

The LBP-TBQ has several strengths compared with previous treatment belief questionnaires. First, our scale addresses several LBP beliefs in relatively compact 16-item or 4-item formats applicable to one or more treatments. This allows a comprehensive assessment with relatively low participation burden compared with existing scales, which assess selected dimensions (eg, overuse and harm in the 8-item BMO).\(^{27}\) Credibility and outcome expectancies in the 6-item CEQ\(^{11}\) and would need to be combined in a longer questionnaire without covering all relevant dimensions (eg, individual fit). Second, previous scales are treatment-specific. For example, deciding on medication use was previously described as involving a cost–benefit analysis comparing perceived benefits with concerns;\(^{26}\) while for other treatments perceived credibility and expected results have been considered more relevant.\(^{12}\) In the LBP-TBQ, identical items assess beliefs about several treatments, facilitating direct comparisons. Third, the LBP-TBQ takes into account acquiescence bias by including negatively worded items, which is rarely considered. For example, medication necessity and concerns have previously emerged as separate dimensions, but distinctions may be partly due to item wording.\(^{27}\) Fourth, using nonparametric item response theory methods allowed us to account for item difficulty and identify latent constructs under less strict (and more realistic) assumptions than parametric item response theory.\(^{28}\) And fifth, exploring MI enabled in-depth understanding of possible sources of variance in questionnaire structure and highlighted areas for improvement.

The LBP-TBQ benefits from combining theory and empirical qualitative data. Theoretically, it conceptualizes patients’ preferences within the CSM as developing from patients’ illness and treatment perceptions\(^{24}\) and therefore assesses patients’ beliefs about treatments to provide information relevant for clinical decision making. According to the CSM (and confirmed in our qualitative research\(^{14}\)), patients need to form adequate illness representations (eg, illness identity and causal representations) to inform their treatment decisions; thus, future studies should assess illness representations alongside treatment beliefs. Using empirical qualitative data, we were able to construct items based on the language LBP patients use to describe their beliefs.

Our validation study also revealed unanticipated and interesting differences between how patients perceive LBP treatments. First, contrary to our hypotheses, we identified significant moderate associations between patients’ concerns and perceptions of individual fit regarding exercise and their concerns about, and emotional representations of, LBP. These associations suggest that unlike other (arguably more passive) treatments, patients with fewer LBP concerns and less intense emotions about LBP have fewer concerns about exercise and stronger beliefs that exercise is right for them. Second, effectiveness beliefs showed the weakest stability, particularly for medication (\(r = 0.63-0.76\)), suggesting that effectiveness beliefs fluctuate more over time than others. It may be that these beliefs are more easily influenced (eg, by personal or vicarious experience of medications or practitioners) than beliefs about credibility, concerns, and individual fit. Understanding the causes and mechanisms of belief variability may be a promising avenue for further research and may reveal appropriate approaches to influencing the development of treatment beliefs. Third, MI was not achieved between treatments, suggesting that patients may interpret items and constructs somewhat differently when assessing different treatments. This is expected given the differences between treatments (eg, one’s concerns regarding medication can be very different to those about exercise) and suggests that future research using the LBP-TBQ for treatment comparisons should first identify the sources of these differences and establish partial invariance.

The MI findings prompt specific recommendations regarding future applications and development of the LBP-TBQ. Scalar invariance was achieved for all treatments between data collection waves and pain duration subgroups, supporting the use of the LBP-TBQ in longitudinal studies, or to compare groups of patients experiencing pain for less or more than 3 years. Not all

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Nerve compression (present or absent)</th>
<th>Sciatica (present or absent)</th>
<th>Pain duration (&lt;3 y or ≥ 3 y)</th>
<th>Treatment experience (present or absent)</th>
<th>Time (1st or 2nd survey)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medication</td>
<td>None</td>
<td>Scalar</td>
<td>Scalar</td>
<td>NA*</td>
<td>Scalar</td>
</tr>
<tr>
<td>Exercise</td>
<td>Scalar</td>
<td>Metric</td>
<td>NA*</td>
<td>Metric</td>
<td>Scalar</td>
</tr>
<tr>
<td>Manual therapy</td>
<td>Metric</td>
<td>Metric</td>
<td>NA*</td>
<td>Scalar</td>
<td>Scalar</td>
</tr>
<tr>
<td>Acupuncture</td>
<td>Scalar</td>
<td>None</td>
<td>NA*</td>
<td>Scalar</td>
<td>Scalar</td>
</tr>
</tbody>
</table>

* Data not available because of low number of treatment-naive patients.

LBP-TBQ, Low Back Pain Treatment Beliefs Questionnaire; MI, measurement invariance.

<table>
<thead>
<tr>
<th>Content</th>
<th>H (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain medication</td>
<td></td>
</tr>
<tr>
<td>Short form</td>
<td>0.43 (0.04)</td>
</tr>
<tr>
<td>Taking/Having [...] for back pain makes a lot of sense</td>
<td>0.58 (0.04)</td>
</tr>
<tr>
<td>I think [...] is pretty useless for people with back pain (d)</td>
<td>0.66 (0.03)</td>
</tr>
<tr>
<td>I have concerns about taking/having [...] for my back pain (r)</td>
<td>0.61 (0.03)</td>
</tr>
<tr>
<td>I am confident [...] would be a suitable treatment for my back pain</td>
<td>0.64 (0.03)</td>
</tr>
</tbody>
</table>

LBP-TBQ, Low Back Pain Treatment Beliefs Questionnaire.

Italic font is used for items with violations of monotonicity, bold font represents significant violations (at default rest group minsize).
treatments displayed scalar invariance between patients with different probabilities of nerve root compression, self-reported sciatica, or treatment experience; patients in these subgroups might attribute different meanings to particular items or constructs. Therefore, the LBP-TBQ structure would benefit from further investigation in relation to different patient characteristics. We recommend researchers using the LBP-TBQ for subgroup comparisons first ascertain whether these subgroups interpret questionnaire items similarly, by performing MI analyses before examining group differences or using cognitive interviews. Moreover, further psychometric work is needed before using the LBP-TBQ to assess beliefs about other treatments (eg, cognitive–behavioural therapy, multidisciplinary rehabilitation), or in other populations, languages, or settings, and we recommend additional qualitative and psychometric examinations of its relevance and item properties in new contexts.

Our research is subject to several limitations. First, our participants were recruited from patients consulting their primary care general practitioner or CAM practitioner; results might be different in secondary care patients (eg, those who are undergoing hospit-based treatments or being considered for spinal injections or surgery). However, some of our patients had experienced hospital-based interventions (eg, 17% had previously attended a pain management clinic, see Supplemental Digital Content 1, available online at http://links.lww.com/PAIN/A79). Second, our participants were older, and more reported unemployment (mostly retired) and chronic pain compared with other primary care LBP samples (eg, mean age: 44 years, 24.6% unemployed [including retirees], 11.4% with pain duration longer than 3 years in another U.K.-based primary care cohort vs 55 years, 52.5% unemployed [including retirees], and 64.8% with pain duration longer than 3 years in our cohort). This is likely due to differences in participant selection (we contacted all LBP patients, not only patients currently consulting their primary care physician) or respondent burden (longer survey in the present study). Thus, the properties of the LBP-TBQ would benefit from confirmation in other samples. Third, although the 4-factor CFA model showed acceptable fit, the subscales were also highly intercorrelated within each treatment and could alternatively be merged into a single “treatment acceptability” dimension. These strong associations could be seen as contrasting with our previous qualitative findings where they emerged as distinct aspects considered when making treatment decisions and may reflect the different sociocognitive processes involved in questionnaire response vs group discussion. Further psychometric work should explore the value of a 4-factor vs 1-factor structure, as well as MI and validity in different populations. Fourth, our study was informed primarily by the NICE care pathway and thus focused on 4 frontline treatments in the United Kingdom; future studies and applications would need to consider local availability and clinical relevance when restricting or expanding this selection. Our study shows that the LBP-TBQ can be used to assess the beliefs of LBP patients regarding 4 guideline-recommended treatment options (pain medication, exercise, manual therapy, acupuncture). According to the extended CSM of illness representations,31 such beliefs are likely to be key modifiable determinants of treatment uptake and adherence. Understanding these beliefs and behaviours in LBP could allow us to develop interventions to optimize the use of evidence-based recommended treatments and thus improve treatment effectiveness and patient outcomes in LBP. Protocols could be developed for clinicians to use the LBP-TBQ to match patients to treatments consistent with their treatment beliefs (eg, Ref. 44) or to identify and address patients’ negative beliefs about treatments. For example, patients could complete the 4-item LBP-TBQ before (eg, electronically) or during consultations (eg, as a clinical interview) regarding locally available treatment options. The patients’ answers could highlight topics for clinicians to explore in the consultation, perhaps after receiving appropriate training;

### Table 8

<table>
<thead>
<tr>
<th>Scale/Item</th>
<th>Mean (SD) α*</th>
<th>Test–retest Mean (SD) α*</th>
<th>Pain medication</th>
<th>Exercise</th>
<th>Manual therapy</th>
<th>Acupuncture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short form</td>
<td>3.58 (0.75)</td>
<td>0.70</td>
<td>0.82</td>
<td>3.60</td>
<td>0.78</td>
<td>0.84</td>
</tr>
<tr>
<td>Taking/Having […] for back pain makes a lot of sense</td>
<td>3.75 (0.91)</td>
<td>0.60</td>
<td>0.74</td>
<td>3.68</td>
<td>0.77</td>
<td>0.80</td>
</tr>
<tr>
<td>I think […] is pretty useless for people with back pain (τ)</td>
<td>3.98 (0.96)</td>
<td>0.67</td>
<td>0.79</td>
<td>3.80</td>
<td>0.63</td>
<td>0.73</td>
</tr>
<tr>
<td>I have concerns about taking/ having […] for my back pain (τ)</td>
<td>3.25 (1.21)</td>
<td>0.71</td>
<td>0.77</td>
<td>3.94</td>
<td>0.54</td>
<td>0.71</td>
</tr>
<tr>
<td>I am confident […] would be a suitable treatment for my back pain</td>
<td>3.34 (0.98)</td>
<td>0.59</td>
<td>0.73</td>
<td>3.45</td>
<td>0.55</td>
<td>0.78</td>
</tr>
</tbody>
</table>

LBP-TBQ, Low Back Pain Beliefs Questionnaire.

* Cronbach alpha.
for example, if a patient reports concerns about the clinically preferred treatment, the clinician could discuss these openly before agreeing on a treatment course. While a less focused discussion of treatment options might be inconclusive, completing the LBP-TBQ could help patients crystallise their beliefs and thus prompt a clinically appropriate decision that reflects patients’ beliefs where possible. Further research is needed to develop and test such clinical applications of the LBP-TBQ.

We recommend the LBP-TBQ for prospective research examining determinants of treatment uptake and adherence in patients with LBP. Because it can assess the same beliefs about different treatments, the LBP-TBQ can also be used to model treatment choices within a more complex decision space than is often considered, i.e., one in which patients are choosing one from among many treatment options. Further research could explore clinical applications of the LBP-TBQ (particularly its 4-item version) to involve patients in treatment decision making and thus help to more systematically take patients’ preferences into account when choosing treatments for LBP.47,48

Conflicts of interest statement
A. Dima declares consultancy work for Baylor Black Sea Foundation Respiratory Effectiveness Group. The remaining authors have no conflicts of interest to declare.

This article presents independent research funded by the National Institute for Health Research (NIHR) School for Primary Care Research (Grant Reference Number 75) and an NIHR Research Professorship for N. E. Foster (NIHR-RP-011-015). F. L. Bishop’s post was funded by Arthritis Research UK (Career Development Fellowship 18099). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

Acknowledgements
The authors thank the participants for their time in completing the survey and for sharing their views, the Primary Care Research Network, and the clinicians who helped to recruit the participants. Cathy Salmon and Sara Oroz assisted in preparing questionnaire packs.

Appendices. Supplemental Digital Content

References
A. Dima et al. • 156 (2015) 1489–1500


[33] Markland D. The golden rule is that there are no golden rules: comment on Markland D. The golden rule is that there are no golden rules: comment on Mao JJ, Xie SX, Bowman MA. Uncovering the expectancy effect: the validation of the Acupuncture Expectancy Scale. Alter Ther Health Med 2010;16:22–7.


[50] Markland D. The golden rule is that there are no golden rules: comment on Markland D. The golden rule is that there are no golden rules: comment on Mao JJ, Xie SX, Bowman MA. Uncovering the expectancy effect: the validation of the Acupuncture Expectancy Scale. Alter Ther Health Med 2010;16:22–7.

[51] Markland D. The golden rule is that there are no golden rules: comment on Markland D. The golden rule is that there are no golden rules: comment on Mao JJ, Xie SX, Bowman MA. Uncovering the expectancy effect: the validation of the Acupuncture Expectancy Scale. Alter Ther Health Med 2010;16:22–7.


