Research Paper

PAIN®

Individual treatment expectations predict clinical outcome after lumbar injections against low back pain

Matthias Müller-Schrader[a,*], Jakob Heinzle[a], Alfred Müller[b], Christian Lanz[b], Oliver Häussler[b], Martin Sutter[b], Andreas Eggspühler[b], Stefania Mare[b], Birte Toussaint[a], Inês Pereira[a], Frederike H. Petzschner[a,c,d], Katja Wiech[e], Jürgen Barth[f], Claudia M. Witt[f], Klaas Enno Stephan[a], Zina-Mary Manjaly[b,g]

Abstract

Subjective expectations are known to be associated with clinical outcomes. However, expectations exist about different aspects of recovery, and few studies have focused on expectations about specific treatments. Here, we present results from a prospective observational study of patients receiving lumbar steroid injections against low back pain (N = 252). Patients completed questionnaires directly before (T_1), directly after (T_2), and 2 weeks after (T_3) the injection. In addition to pain intensity, we assessed expectations (and certainty therein) about treatment effects, using both numerical rating scale (NRS) and the Expectation for Treatment Scale (ETS). Regression models were used to explain (within-sample) treatment outcome (pain intensity at T_3) based on pain levels, expectations, and certainty at T_1 and T_2. Using cross-validation, we examined the models' ability to predict (out-of-sample) treatment outcome. Pain intensity significantly decreased ($P < 10^{-15}$) 2 weeks after injections, with a reduction of the median NRS score from 6 to 3. Numerical Rating Scale measures of pain, expectation, and certainty from T_1 jointly explained treatment outcome ($P < 10^{-15}$, $R^2 = 0.31$). Expectations at T_1 explained outcome on its own ($P < 10^{-10}$, $f^2 = 0.19$) and enabled out-of-sample predictions about outcome ($P < 10^{-4}$), with a median error of 1.36 on a 0 to 10 NRS. Including measures from T_2 did not significantly improve models. Using the ETS as an alternative measurement of treatment expectations (sensitivity analysis) gave consistent results. Our results demonstrate that treatment expectations play an important role for clinical outcome after lumbar injections and may represent targets for concomitant cognitive interventions. Predicting outcomes based on simple questionnaires might be useful to support treatment selection.

Keywords: Chronic pain, Low back pain, Expectation, Belief, Regression, Out-of-sample prediction, Longitudinal study

1. Introduction

In many areas of medicine, it is well established that outcomes of clinical interventions can be shaped significantly by individual expectations.[2,27] In particular, concerning the clinical management of pain, expectation effects on the experience of pain have not only been demonstrated by many experimental studies, eg, in studies on placebo and nocebo responses,[6,17,23,38,43,45] but have also been found in a variety of clinical settings.[5,9,12,18,21]

Importantly, in clinical settings, patients can have expectations along multiple dimensions.[32] For example, expectations may exist simultaneously about recovery in general (unrelated to a specific treatment), about one's ability to cope with pain (self-efficacy), and about the outcome of a specific treatment. A recent Cochrane review on low back pain[19] evaluated the evidence for the impact of different types of expectations. It highlighted that many investigations of general expectations about recovery exist, whereas little is known about the impact of treatment-specific outcome expectations.

Here, we report results from a prospective observational study that addresses this gap. We investigated 252 patients with low back pain who underwent lumbar injections of steroids and completed several questionnaires directly before (T_1), directly after (T_2), and 2 weeks after (T_3) the intervention. Injection therapy is a useful setting for investigating expectation effects because of the controlled nature of the therapeutic intervention and lack of potential problems of patient compliance. Additionally, injections are brief, salient interventions that allow for anchoring treatment expectations to a well-defined point in time.

Our study is novel in 2 additional ways. First, our study was inspired by contemporary "Bayesian brain" theories of perception in general (eg, "predictive coding" or "predictive processing"[15,24]) and of pain experience in particular.[8,16,20,31,40] Consequently, we did not only measure expectations about treatment outcome but

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

[a] *Translational Neuromodeling Unit (TNU), Institute for Biomedical Engineering, University of Zurich & ETH Zurich, Zürich, Switzerland,* [b] *Department of Neurology, Schulthess Clinic, Zurich, Switzerland,* [c] *Department of Psychiatry and Human Behavior, Brown University, Providence, RI, United States,* [d] *Robert J. and Nancy D. Carney Institute for Brain Science, Brown University, Providence, RI, United States,* [e] *Wellcome Centre for Integrative Neuroimaging (WIN), Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe Hospital, Oxford, United Kingdom,* [f] *Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Zurich, Switzerland,* [g] *Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland*

**Corresponding author. Address: Translational Neuromodeling Unit (TNU), University of Zurich & ETH Zurich, Wilfriedstrasse 6, 8032 Zürich, Switzerland. Tel.: +41 44 634 91 09. E-mail address: mueller-schrader@biomed.ee.ethz.ch (M. Müller-Schrader).*

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

PAIN 00 (2022) 1–10

http://dx.doi.org/10.1097/j.pain.0000000000002674

also their "precision," ie, how certain patients were in their own expectations about treatment outcome. Together, expectations and precision (or certainty) characterise patients' beliefs about treatment outcomes more comprehensively than classical measures of expectations alone. In addition, considering the potential effects of belief precision (certainty) on treatment outcomes is motivated by recent work on placebo effects[1,17,22] and theories of bodily regulation,[39] which suggest that the precision of treatment beliefs contributes to therapeutic outcomes, over and beyond expectations.

Second, in addition to conventional statistical analyses that determine statistical associations within-sample, we asked whether it was possible to *predict* (out-of-sample) future pain levels from expectations. More specifically, we examined whether reported pain intensities 2 weeks after the injection (at T_3) could be predicted from expectations and/or certainty before treatment (T_1). Addressing this question is clinically important: a procedure to predict the outcome of specific interventions based on simple questionnaires about expectations and/or certainty might not only support individual treatment selection, but would also identify these variables as targets for concomitant cognitive interventions to boost the efficacy of injection therapy (for expectation-focused interventions in other domains, see [11,27,33]).

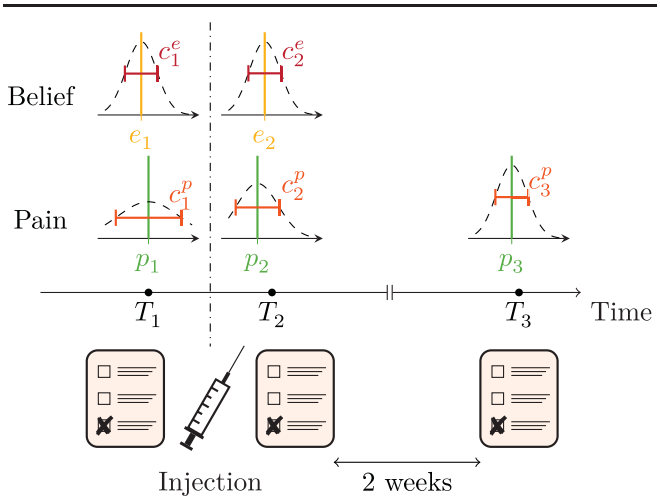

Figure 1. Visual summary of the study design. Pain ratings and measures of beliefs (about treatment outcome) were obtained at 3 different time points T_1 to T_3 (at T_3 pain ratings only, which served as outcome measure). At each time point T_i, patients were asked to indicate pain level (p_i) and expectation about treatment outcome (e_i) and confidence therein (c_i^p/ c_i^e).

2. Methods

2.1. Design and participants

Three hundred six patients with low back pain participated in this study which employed an observational longitudinal design in a naturalistic setting. All patients received the same treatment, ie, lumbar injections under radiological guidance, at the Department of Neurology, Schulthess Clinic Zurich. The injections included a local anesthetic (either bupivacaine 0.5% or lidocaine 0.5%) and a steroid (either triamcinolone acetonide extended release 40 mg or 80 mg or betamethasone 17α,21 dipropionate 5 mg + betamethasone 21-disodium phosphate 2 mg).

Pain ratings and measures of belief (about treatment outcome) were obtained at 3 different times: directly before (T_1) and immediately after (T_2) the injection, as well as 2 weeks after the injection (T_3). Questionnaires for T_3 were sent and answered by mail. The exact wording of the questions and the definition of the rating scales are provided in the Supplementary Material (available at http://links.lww.com/PAIN/B636).

At each time point *i*, we obtained 2 main measures (see **Fig. 1** for a summary). First, we asked participants to indicate their perceived level of pain intensity (p_i) on a Numerical Rating Scale (NRS), ranging from 0 to 10. Using a Visual Analogue Scale (VAS), we also asked them how certain they were about the pain intensity they indicated (c_i^p).

Second, we asked participants about their subjective beliefs about treatment outcome. Here, we distinguished between the expectation about treatment outcome e_i (ie, the intensity of pain they expected to have 2 weeks after the treatment) and the certainty c_i^e in this expectation. Expectations were assessed using a NRS (0-10), and certainty was measured using a VAS, in the same way as for certainty about pain.

In addition to the VAS, we tried to obtain an additional, novel measure of certainty. Using a NRS (0-10), we asked participants to exclude those levels of pain or expectation, respectively, which they could exclude with certainty. However, the responses clearly indicated that our participants found this novel measure confusing: More than half of the participants gave responses which were inconsistent with their statements on their current pain intensity and/or expectations. We therefore decided not to use this measure of certainty in our analyses.

Since different measures of treatment expectations have been proposed in the past, we intended to perform a sensitivity analysis and examine whether our results were robust, independent of this choice. For this reason, we obtained an alternative measure of subjective treatment expectations, using the Expectation for Treatment Scale (ETS[4]). This questionnaire represents a useful alternative measure for a sensitivity analysis because, in contrast to our NRS described above, it mainly asks about expected improvements, less about the expected outcome per se. Complete ETS data were obtained from 246 of the 252 participants included in our analysis. Data from 6 participants were missing because these participants did not manage to complete all questionnaires in the available time before treatment started.

Finally, at T_1, we assessed several variables that might explain variability of individual beliefs about treatment outcome. Specifically, these variables included previous injection therapy and its success, educational background, pain-specific self-efficacy,[30] relevance of the opinions of others for the decision to undergo injection therapy, and duration of treatment-requiring pain. Please see the Supplementary Material (available at http://links.lww.com/PAIN/B636) for details how these measures were obtained. Complete measures of these variables were available for 183 of the 252 participants included in our analysis. Data from the remaining participants were missing because these participants did not manage to complete all questionnaires prior to treatment.

This study received a waiver from the Cantonal Ethics Committee (Req-201900332). Nevertheless, we obtained written informed consent from all participants before they participated in this study. Participation in this study did not have any influence on the treatment. Inclusion criteria were broad (ie, any patient receiving lumbar injections against low back pain in our clinical department), and the only exclusion criterion was an inability to fill in the required questionnaires (eg, due to acute pain or insufficient command of the German language).

Our sample size was guided by an a priori power analysis. Given the lack of previous data, the power analysis concerned the

question whether future pain intensities were explainable by any subset of the variables (ie, an F-test on all regressors in the most comprehensive model m_2; for its definition, see below). We adopted a conservative approach, assuming a small effect size (ie, $f^2 = 0.1$) and requiring a power of 0.9, at a significance level of 0.05. Under these assumptions, the sample size needed was N = 215. Given concerns about attrition and incomplete questionnaires, we decided to add a substantial safety margin, under the anticipation that data from approximately 1/3 of the participants might not be usable. This assumption turned out to be too pessimistic: Data from 54 of 306 (17.6%) participants could not be used. Therefore, data of 252 patients (52% female) were included in our main analysis. The median age was 70 years (minimum 19 years and maximum 93 years). Note that the detectable effect size for some of the analyses is considerably lower than the $f^2 = 0.1$ used in the power analysis. The smallest detectable partial effect of confidence c_1^e in the smaller model m_1 has for example an effect size of $f^2 = 0.04$ (with power and significance level as above).

2.2. Statistical analyses

As described in the Introduction, our study was inspired by general "Bayesian brain" concepts of perception.[15,24] One postulate of these concepts is that perception corresponds to a posterior belief (about the state of the world) that results from integrating sensory data (in statistical terms: the "likelihood") with prior beliefs. Notably, these prior beliefs are not only characterised by expectations; instead, their precision (ie, subjective certainty) matters as well. This motivated our experimental strategy (as described above) to collect measures of both subjective expectations about treatment outcome and the subjective certainty therein.

Given this background, a natural analysis strategy would have been to construct a fully Bayesian model of how posterior beliefs result from updating prior beliefs. However, the clinical setting of our study imposed constraints on the type of data we could acquire, leaving only time for limited questionnaire-based assessments. Given these available data, constructing a fully Bayesian model faces major challenges. These challenges required us to resort to an approximation, as explained in the following.

The most critical problem concerns the question how one would obtain an estimate of the likelihood. For example, this would require experimental control over the mean and variance of the sensory inputs that result from the cause of pain (eg, pain-inducing stimuli). In our clinical setting, this was not possible, since the dense clinical routine only allowed for a brief assessment using questionnaire-based self-reports. This makes the construction of a model which describes the transition from prior to posterior difficult.

However, even in the absence of precise information about the likelihood, we know that (all else being equal) a variation in the prior precision should lead to a variation in the posterior mean (and posterior precision). In other words, provided one has valid measures of subjective expectations and confidence, one would expect variations in these statistics of prior beliefs to result in variations of posterior mean and precision. In the absence of knowledge about the exact mathematical form of this relationship, a linear model can be chosen as the simplest approximation. Clearly, this approximation may fail if the relationship is sufficiently non-linear. The practical advantage of this approximation is, however, that it allows for constructing multivariate regression models that explain pain intensity after treatment as a linear combination of the questionnaire-based measures of subjective expectations and certainty. Below, we describe several of these models that examine different hypotheses.

Generally, it should be kept in mind that, while our analyses are inspired by a Bayesian perspective on perception, they do not employ a proper Bayesian model and should therefore not be viewed as trying to confirm or disprove a Bayesian perspective on pain experience.

The analysis tested 7 hypotheses which were specified ex ante in a time-stamped analysis plan (https://gitlab.ethz.ch/tnu/analysis-plans/muellerschraderetal_pain_expectation_2021). Bonferroni correction resulted in a corrected significance level of $\overline{\alpha} = 0.05/7 = 0.0071$. For comparisons of pain levels across time points, we report Cohen's d as effect size. In the context of multiple regression models, we used R^2 to report the global effect size (ie, for the entire regression model) and Cohen's f^2 to report local effect sizes (ie, for variables of interest). All effect sizes are reported with 90% confidence intervals. The analysis was conducted using the open-source statistical computing language R, version 4.1.2,[37] and several of its packages.

2.3. Regression analysis

For our analyses, we constructed several multiple linear regression models which explain treatment outcome, ie, pain intensity p_3 at T_3, as a function of variables assessed at T_1 and, in some cases, also T_2 (compare **Fig. 1**).

The simplest model, m_0, included only 3 regressors: the initial pain level p_1 as an explanatory variable and age a and gender g as confounds. Hence, m_0 constituted a *basic model* to explain treatment outcome p_3 without reference to beliefs. All other models were extensions of this model, with additional regressors that represented measures of beliefs (expectations and certainties). We used F-tests to assess whether the inclusion of additional regressors significantly improved models.

Model m_1 contained all main measures obtained before the injection (ie, at T_1): in addition to p_1, a, and g, we included the expectation about treatment outcome, e_1, and the certainty in this expectation, c_1^e. Furthermore, we also included c_1^p, that is, how certain they were about the indicated pain intensity p_1.

We assessed (F-test) whether m_1 explained the treatment outcome significantly better than m_0. Two subsequent F-tests focused on the specific roles of treatment expectation e_1 and certainty c_1^e within m_1. First, we tested whether the subjective belief in toto (ie, the combined effect of expectation and certainty) contributed significantly to explaining treatment outcome p_3. Second, we tested whether treatment expectation and certainty, respectively, showed significant relations to treatment outcome on their own.

2.4. Extensions of model

Next, we tested 2 extensions of m_1: model m_{1+int}, which additionally included interaction terms (see below), and m_2, which extended m_1 by including measures obtained at T_2 (directly after the injection).

Model m_{1+int} included the interaction terms $p_1 \cdot c_1^p$ and $e_1 \cdot c_1^e$ as well as $e_1 \cdot p_1$. The inclusion of these interaction terms is motivated by the Bayesian perspective on pain experience described in the Introduction. In brief, under Gaussian assumptions, Bayesian treatments of perception describe the updating of beliefs as a precision-weighted compromise between sensory data and prior expectations.[35] Here, we attempt a crude approximation to this

principle by considering the interaction of certainty (as a measure of precision) with pain and expectation, respectively.

Model m_2 included measures of pain and beliefs obtained directly (within 30 minutes) after the injection (T_2). Placebo research suggests that the invasiveness of a procedure is related to the strength of placebo effects.[32] Hence, individual differences in the experienced invasiveness of the injection might induce variations in expectations about longer-term outcome. Similarly, immediate therapeutic effects (due to the anesthetic that is administered together with steroids) that change experienced pain intensity right after injection can occur. For model m_2, we therefore included the additional regressors e_2, c_2^e, p_2, and c_2^p.

For both models, we used an F-test to assess whether the additional regressors (compared to m_1) significantly improved explanation of treatment outcome p_3.

2.5. Out-of-sample prediction

After checking whether the above models can *explain* the treatment outcome (within-sample), we examined whether it was also possible to *predict* the treatment outcome (ie, pain intensity p_3 at T_3) out-of-sample, based on data acquired before (T_1) or shortly after (T_2) the injection. We conducted this analysis using the models m_1, m_2, and m_{1+int} described above. As a comparison, we also included the basic model m_0.

For prediction, we employed a 10-fold cross-validation scheme. In each fold k, we used the training data (9/10 of the data) to estimate the regression coefficients β of our models. We then tested the prediction $\hat{p}_3^n$ on the held-out sample, M_k, using the estimates of β from the training data. Here, $n \in M_k$ indexes the held-out samples, and p_3^n denotes the treatment outcome (pain intensity at T_3) for participant n.

To assess the goodness of the prediction, we used 2 separate metrics. For each fold, we calculated these metrics over all held-out participants in that fold and averaged these metrics over all folds. First, we used the root mean squared error (RMSE):

$$RMSE_k = \sqrt{\frac{1}{|M_k|}\sum_{n\in M_k}\left(\hat{p}_3^n - p_3^n\right)^2} \quad (1)$$

Here, $|M_k|$ denotes the cardinality of the fold (ie, how many participants were held out).

As a second metric, we used the coefficient of determination (R^2):

$$R_k^2 = 1 - \frac{\sum_{n\in M_k}\left(\hat{p}_3^n - p_3^n\right)^2}{\sum_{n\in M_k}\left(\hat{p}_3^n - \bar{p}_3\right)^2} \quad (2)$$

Here, $\bar{p}_3$ denotes the mean of p_3.

For each metric, the cross-validation procedure resulted in one value per model. We used permutation testing to statistically assess the significance of these results. We randomly chose 10,000 permutations of the numbers from 1 to N, where N denotes the number of participants. For each of these permutations π, we permuted the treatment outcomes p_3^n (keeping the rest of the data fixed) so that we arrived at new tuples $(p_1^n, e_1^n, \cdots, p_3^{\pi(n)})$ (for $n = 1, ..., N$). This procedure completely removes any dependence between p_3 (the outcome) and the predictors. We then performed the same cross-validation procedure for all permutations, yielding 10,000 values of our metrics per model, which represent predictions under the null hypothesis that there is no relation between the predictors and the outcome. P values were calculated with reference to this null distribution.

So far, we examined how well the models m_1, m_2, and m_{1+int} were able to predict *relative to chance*. To analyse how capable the models were of predicting *over and beyond the basic model* m_0, we defined $\tilde{p}_3$, the residuals when explaining p_3 by m_0 in the training data set. We repeated the analyses (cross-validation and permutation testing) with $\tilde{p}_3$ as outcome. If any model was still able to predict better than chance, this prediction would have to rely on information not captured by model m_0.

2.6. Formation of beliefs

In this analysis, we investigated which of 5 candidate variables (previous injection therapy, educational background, pain-specific self-efficacy, weight of the opinions of others for choosing injection therapy, and duration of treatment-requiring pain) might explain variability of individual beliefs about treatment outcome before the treatment was administered (ie, at T_1). To this end, we constructed 2 separate multiple regression models that explained treatment expectation e_1 and certainty c_1^e, respectively, using the above variables together with measures of pain (p_1 and c_1^p) as regressors. Using F-tests, we tested, for each of the variables of interest, whether it was significantly associated with treatment expectation e_1 or certainty c_1^e, respectively.

The question whether any of our candidate variables bears a relation to beliefs constitutes 1 of the 7 hypotheses tested in this study (see above and the analysis plan for details). This particular question, however, necessitates 10 distinct tests. To avoid inflating our overall false-positive rate, for this particular question, we used the Benjamini–Hochberg FDR-correction with a false discovery rate equal to the overall Bonferroni-corrected level of $\overline{\alpha} = 0.0071$.

2.7. Sensitivity analysis: replacing beliefs with Expectation for Treatment Scale scores

The analyses above showed that the belief at T_1 allowed for explaining (within-sample) and predicting (out-of-sample) the treatment outcome 2 weeks later, at T_3. In order to check the robustness of these results and to test whether they would change qualitatively if a different construct were used, we replaced the measured components of subjective beliefs (treatment outcome expectation e_1 and certainty c_1^e). For this sensitivity analysis, we used the measure of treatment outcome expectation provided by the Expectation for Treatment Scale (ETS[4]) at T_1.

This resulted in a new model $m_{1,E}$, which had the same regressors as m_1, except that e_1 and c_1^e were replaced by the (summed) ETS score. Using an F-test, we assessed whether the ETS was significantly associated with treatment outcome at T_3. Furthermore, we performed out-of-sample predictions, as for the other models.

3. Results

In this section, we first present the raw data as histograms. We then show the results of the regression analyses and finally proceed to the out-of-sample prediction of pain at outcome (T_3).

3.1. Histograms of data

Figure 2 shows the distribution of the main data obtained directly before (T_1), directly after (T_2), and 2 weeks after (T_3) the injection. Using a two-sided Wilcoxon signed-rank test with continuity

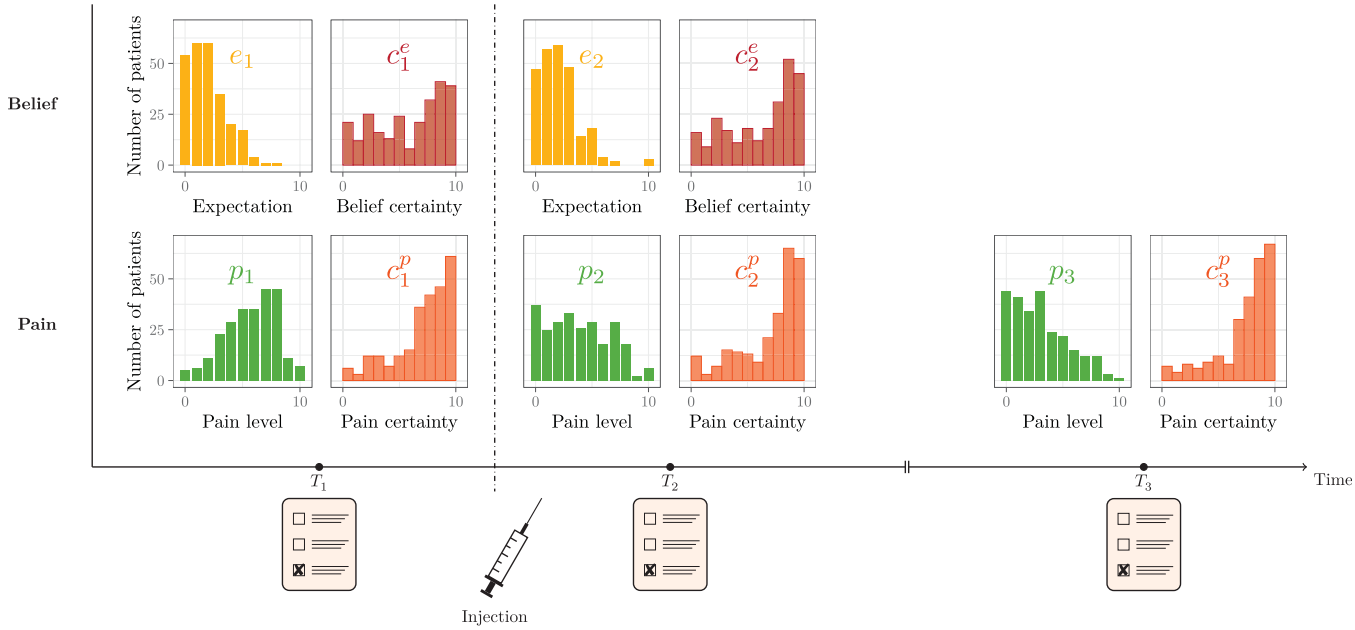

Figure 2. Presentation of raw data. Note that expectation and pain levels are discrete, while the certainty measures are continuous. Variables indicate pain levels (p), expectations (e), or certainty (c); subscripts indicate time (please see Fig. 1); superscripts indicate whether certainty relates to pain (p) or expectation (e).

correction, we found that compared to T_1, pain intensity was significantly reduced ($P = 2.2{\cdot}10^{-16}$; $V = 22654$) 2 weeks after treatment at T_3 ($median(p_1) = 6$, $median(p_3) = 3$). Between T_1 and T_2, there was no significant change in treatment outcome expectations e_i ($P = 0.13$; $V = 2520.5$) and in certainty about pain intensity c_i^p ($P = 0.69$; $V = 11880$), while there was a significant change in pain intensity p_i ($P < 2.2{\cdot}10^{-16}$; $V = 14337$) and certainty about treatment outcome expectation c_i^e ($P = 0.015$; $V = 9580$).

3.2. Regression analysis

The *basic model* m_0 already explained a significant amount of the variability of treatment outcome p_3 ($F_{3,248} = 16.23$; $P = 1.2 \cdot 10^{-9}$; $R^2 = 0.16$ [0.10, 0.23]). Including measures of beliefs- ie expectation of treatment outcome e_1 and certainty therein, c_1^e – improved the model significantly: in m_1, the expectation e_1 and certainty c_1^e jointly contributed significantly ($F_{2,245} = 27.023$; $P = 2.5 \cdot 10^{-11}$; $f^2 = 0.22$ [0.12, 0.33]) to the treatment outcome, showing that beliefs had a strong relation to treatment outcome over and beyond the basic information (pain intensity, age, and gender) provided by m_0. When examining expectations and certainties separately, we found a significant contribution by expectation e_1 on its own ($F_{1,245} = 47.205$; $P = 5.3 \cdot 10^{-11}$; $f^2 = 0.19$ [0.11, 0.30]), but not by certainty c_1^e ($F_{1,245} = 1.5649$; $P = 0.21$; $f^2 = 0.007$ [0.00, 0.03]). Taking into account all regressors in m_1, **Figure 3d** shows the relation between expectation e_1 and treatment outcome p_3: The partial correlation is 0.40 ($P < 5.3 \cdot 10^{-11}$, $t = 6.87$) and the partial regression coefficient is $\beta = 0.59$. This means that a change of expectation by 1 point corresponds to a change of 0.59 in p_3 (treatment outcome) on the 0 to 10 NRS (when adjusting for the other factors).

3.3. Extensions of model by interaction terms and data from T_2

Neither extending model m_1 with interaction terms (m_{int}: $F_{3,242} = 0.3882$; $P = 0.76$; $f^2 = 0.005$ [0.00 0.02]) nor extending it with data from T_2 (m_2: $F_{4,241} = 1.4164$; $P = 0.23$; $f^2 = 0.02$ [0.00 0.05]) improved explanation of treatment outcome p_3 significantly.

3.4. Out-of-sample prediction

After having established a relation of expectations with treatment outcome in the above regression analyses, we attempted to predict treatment outcome out-of-sample. To this end, we employed the same regression models as above but used a 10-fold cross-validation procedure to train the model on part of the data (training data) and predict treatment outcomes for the held-out participants.

The results are summarized in **Figure 4**. The first thing to note is that all models yield predictions that are better than chance. More precisely, for all models (m_0, m_1, m_{int}, and m_2) and all metrics (R^2 and RMSE), compared with null distributions consisting of 10,000 random permutations, at most 1 sample showed higher accuracy than the real data. This suggests that the result is unlikely to occur by chance ($P = 10^{-4}$). Furthermore, it is worth noting that the most predictive model (ie, the model showing the highest R^2 and lowest RMSE; **Fig. 4A and B**) was model m_1, that is, the model including beliefs about treatment outcome before the injection took place (T_1). This model showed a median absolute deviation (MAD) of 1.32. From a practical perspective of clinical utility, this finding means that the median error of prediction is 1.32 when using a Numerical Pain Rating Scale with a range from 0 to 10.

Next, we investigated the improvement in prediction that was afforded by including beliefs about treatment outcome. First, we examined whether the prediction improvement achieved by models m_1, m_{int}, and m_2 (which included beliefs about treatment outcome) relative to our basic model m_0 (which did not include beliefs) was statistically significant. We calculated P values for predictions over and beyond the null model m_0 (compare also **Fig. 4C and D**). These were smaller than $6{\cdot}10^{-4}$ for all models (m_1, m_{int}, and m_2) and both metrics (R^2 and RMSE), indicating that the inclusion of beliefs about treatment outcome significantly improved predictions.

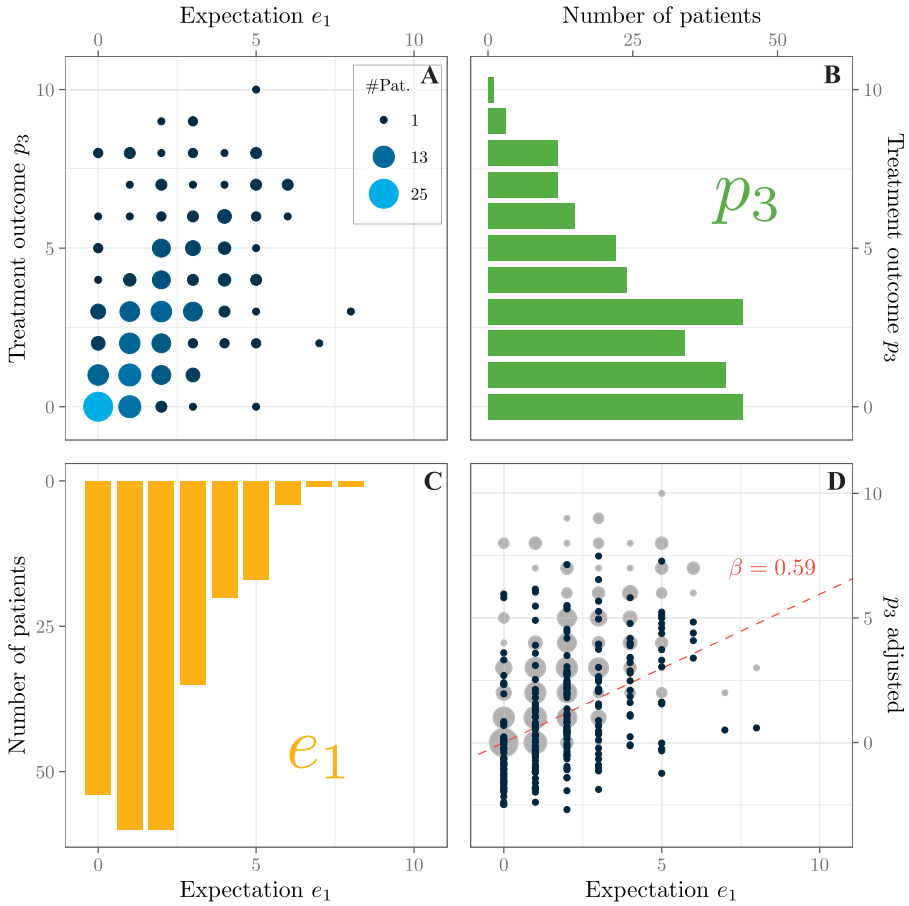

Figure 3. (A) Relation of initial expectation e_1 and treatment outcome p_3 illustrated as a distribution of the number of patients per combination of e_1 and p_3. (B/C) Marginal distributions of p_3 and e_1, respectively. (D) Relation between the expectation e_1 and the adjusted treatment outcome p_3 (ie, that part of the treatment outcome which cannot be accounted for by other regressors in m_1). The gray dots show the full treatment outcome p_3 (as in A). The red dashed line indicates the partial regression of e_1 in m_1. e_1 denotes the expectation at time T_1; p_3 indicates treatment outcome (pain level) at time T_3.

In a second step, we quantified the degree of improvement that was afforded by including beliefs about treatment outcome. For this purpose, we focused on the most predictive model (m_1) and the basic model (m_0) and examined the differences in R^2 and RMSE compared with the null distributions generated by the permutations. We found that the RMSE was reduced by approximately 8% using m_0, but roughly 17% when using m_1. Furthermore, R^2 increased from 0.18 when using m_0 to 0.30 for m_1.

3.5. Formation of beliefs

The analyses described above demonstrated that subjective beliefs about treatment outcome prior to therapy (at T_1) were significantly associated with and strongly contributed to predicting the actual outcome. In a subsequent analysis, we investigated which variables might have shaped individual beliefs (ie, both expectations and certainty therein) about treatment outcome. We focused on the following 5 candidate factors: previous injection therapy, educational background, pain-specific self-efficacy, opinions of others for the decision to undergo injection therapy, and duration of treatment-requiring pain. Using multiple regression, we tested which of these variables were significantly associated with treatment expectation e_1 or certainty c_1^e, respectively.

The overall regression model explained a significant amount of variance in both subjective expectation e_1 ($P = 2 \cdot 10^{-5}$; $F_{16,166} = 3.51, R^2 = 0.25$ [0.17, 0.33]) and certainty c_1^e ($P = 7.8 \cdot 10^{-5}$; $F_{16,166} = 3.21$, $R^2 = 0.24$ [0.16, 0.32]) at T_1. When investigating the specific influence of each of the 5 candidate variables on expectation and certainty, only 1 of the 10 pairs showed a significant effect after correcting for multiple comparisons (compare **Table 1**). Specifically, the opinions of others concerning injection therapy contributed significantly to explaining certainty c_1^e ($F_{1,166} = 13.439$; $P = 3.3 \cdot 10^{-4}$; $f^2 = 0.07$ [0.02 0.15]).

Originally, we had intended to perform a mediation analysis to examine whether a possible influence of these additional measures on treatment outcome p_3 was mediated by beliefs, but since we did not find a significant effect of c_1^e on p_3, there was no possible causal path to perform the mediation analysis on.

3.6. Sensitivity analysis: replacing measures of beliefs with Expectation of Treatment Scale scores

In order to verify that our results concerning the influence of expectations on treatment outcome did not depend on the particular choice of measurement scale (a standard NRS ranging from 0 to 10), we conducted a sensitivity analysis. For this purpose, we chose the ETS as an alternative measure of expectations about treatment outcome. ETS measures were available for 246 participants; the tau-equivalent reliability (Cronbach's alpha) of these measures was very good (0.85; 95% CI: 0.82-0.88).

As expected, we found that ETS scores were negatively correlated with e_1 ($\rho = -0.49, t = -8.67, P = 6.3 \cdot 10^{-16}$); they

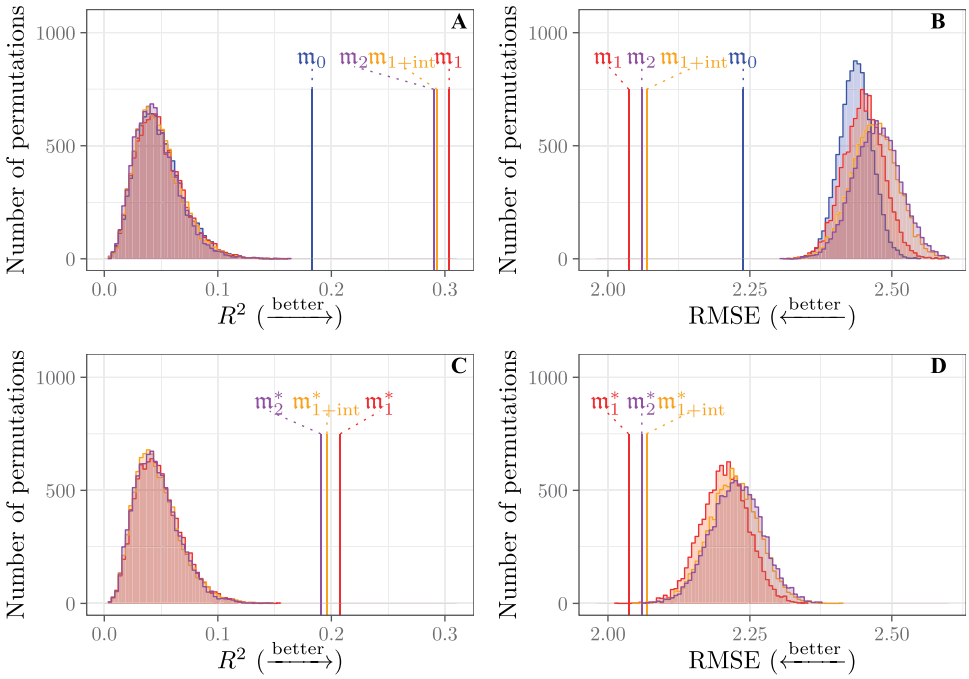

Figure 4. Prediction of treatment outcome p_3: Cross-validated R^2 (A/C) and RMSE (B/D). The histograms represent null distributions and are created by permuting the labels of the treatment outcome p_3. (A) and (B) show the predictive performance on the treatment outcome p_3. (C) and (D) illustrate the results on the residuals $\tilde{p}_3$ after using m_0 to explain the data (also indicated by "*" in the model names).

were also positively correlated with c_1^e ($\rho = 0.18$, $t = 2.86$, $P = 0.0045$).

Using the ETS score instead of the belief measures e_1 and c_1^e as part of the regression model m_1 led to qualitatively similar results: ETS scores also showed a significant ($F_{1,240} = 20.503$; $P = 9.4 \cdot 10^{-6}$; $f^2 = 0.08$ [0.03, 0.16]) relation to treatment outcome (model $m_{1,E}$). **Figure 5** (which is structured analogously to **Fig. 3**) visualises the relation between ETS and treatment outcome.

Similarly, the out-of-sample analyses were able to predict treatment outcome p_3 at T_3 from measurements obtained before the injection at T_1 (R^2: $P < 10^{-4}$, RMSE: $p < 10^{-4}$). Prediction was also still possible on the residuals after having applied m_0 (R^2: $P = 0.0216$, RMSE: $P = 0.0039$).

4. Discussion

This study on the impact of treatment expectation effects on the outcome of injection therapy in low back pain had a triple motivation.

First, as highlighted by a recent systematic review,[19] numerous studies have investigated effects of general recovery expectations on low back pain. By contrast, the role of expectations about specific treatments has received far less attention. Here, we focused on injection therapy because it is frequently used, is not affected by potential compliance problems, and allows for anchoring treatment expectations to a well-defined time point. We are not aware of any previous study examining the impact of expectations on outcomes of injection therapy.

A second aim of our study was to test predictions from the "Bayesian brain" theory of perception (also referred to as "predictive coding" or "predictive processing"[15,24]). This concept has been remarkably successful in explaining a wide range of perceptual phenomena[25,34] and is increasingly applied to explain the experience of pain.[8,16,20,31,40] In brief, the "Bayesian brain" theory holds that the brain represents states of the world (including the body) in terms of "beliefs" (probability distributions). These beliefs are characterised, at a minimum, by an expectation (mean) and certainty (or precision, the inverse of variance). Furthermore, the Bayesian perspective implies that pain perception results from integrating prior beliefs and sensory inputs, weighted by their relative precision. In other words, expectations exert stronger effects if imbued with certainty (ie, if the belief is precise). This implies that studies of expectation effects on pain would benefit from incorporating measures of certainty or precision. Indeed, recent studies on placebo analgesia[1,17,22] have provided evidence of the importance of belief precision for pain perception. However, to our knowledge, it has not been tested whether measures of belief precision also contribute to explaining and predicting treatment outcomes.

A third goal of our study was to test whether individual responses to injection therapy could be predicted from individual expectations and certainties prior to therapy. This analysis adopted cross-validation to investigate whether out-of-sample predictions, as opposed to within-sample statistical associations, would be possible. A positive answer to this question would have clinical implications: The ability to predict individual treatment response from expectations and certainties could guide treatment selection and might offer novel targets for boosting existing therapies by cognitive interventions.[11,27,33] While some attempts of predicting future pain states exist,[3,7,28,29,41,44] these are mostly not tied to a specific intervention. Furthermore, to our knowledge, there is no study that has tried to exploit individual expectations for predicting treatment outcomes out-of-sample. Although implied by the wording in some papers,[9,26] previous studies on the relation between expectations and treatment outcomes have not presented out-of-sample predictions, but within-sample analyses of statistical associations.

Our multiple regression analyses examined which pretreatment variables (T_1) explained treatment outcome after 2 weeks (T_3). In addition to our variables of interest (treatment expectation

Table 1

Results of testing whether any of the candidate variables could explain beliefs about treatment outcome (consisting of expectation e_1 and certainty c_1^e) at T_1.

	Explaining expectation e_1		Explaining certainty c_1^e	
	P	f^2	P	f^2
Previous injection therapy	0.03	0.03 [0.00, 0.09]	0.66	0.004 [0.00, 0.02]
Educational background	0.18	0.03 [0.00, 0.07]	0.61	0.02 [0.00, 0.03]
Pain-specific self-efficacy	0.04	0.04 [0.01, 0.11]	0.01	0.02 [0.00, 0.07]
Opinions of others	0.34	0.003 [0.00, 0.03]	$3.31 \cdot 10^{-4}$*	0.07 [0.02, 0.15]
Duration of strong pain	0.57	0.008 [0.00, 0.03]	0.97	0.001 [0.00, 0.00]

* *P* values reported in this table are uncorrected. When corrected for multiple comparisons, only the weight of the opinions of others in the decision to undergo injection therapy made a significant contribution to explaining certainty c_1^e.

and certainty), our models also included baseline status (pain intensity at T_1) and potential confounds (age and gender). We found that beliefs about treatment outcome at T_1 were highly significantly associated with the actual treatment outcome at T_3. When examining the belief components separately, we found a significant association for expectations, but failed to detect such an association for certainty (belief precision). It is worth emphasising that the latter finding should not be misunderstood as disproving the "Bayesian brain" perspective on pain perception. This is for 2 reasons. First, the validity of our questionnaire-based measures of subjective certainty is unclear: There is presently no validated questionnaire of this sort; additionally, one of our questionnaires had to be discarded due to logical inconsistencies in participants' responses. Second, given methodological problems of constructing a fully Bayesian model (see Methods), we used a simple linear approximation. This approximation may have failed to adequately capture the nonlinear relationship between prior and posterior beliefs. For these reasons, the relevance of belief precision (certainty) for explaining treatment outcomes remains an open question that should be revisited with improved questionnaires and models.

Using the same multiple regression models for out-of-sample predictions (10-fold cross-validation) showed that utilising

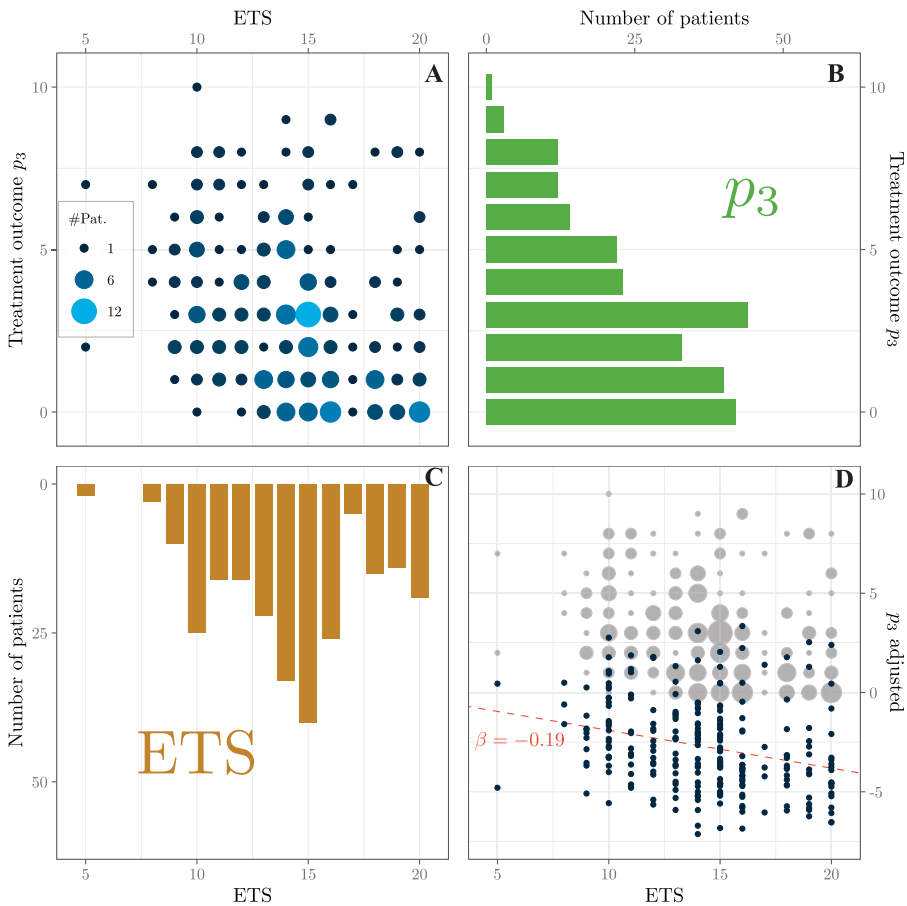

Figure 5. (A) Relation of ETS and treatment outcome p_3 illustrated as a distribution of the number of patients per combination of ETS and p_3. (B/C) Marginal distribution of ETS and p_3, respectively. (D) Relation between ETS and the part of the treatment outcome p_3 that has not been accounted for by other regressors in $m_{1,E}$. The gray dots show the full treatment outcome p_3 (as in A). The red dashed line indicates the partial regression of ETS in $m_{1,E}$. ETS, Expectation for Treatment Scale.

information about individual beliefs (expectations and certainty) considerably improved predictions: Incorporating pretreatment beliefs at T_1 allowed for highly significant predictions about treatment outcome at T_3, with a median error of 1.36 (on a NRS of 0-10).

These results demonstrate the importance of individual expectations for treatment outcomes in patients receiving lumbar injections against low back pain. Our findings have 2 important clinical implications. First, they suggest that success of injection therapy could be enhanced by cognitive interventions that optimize expectations (compare[11,27,33]); the model suggests a reduction of pain of 0.59 per point improved expectations (on 0-10 NRSs). Additionally, the demonstration that it is possible to predict individual treatment response based on information obtained through simple questionnaires might support treatment selection, for example, determining whether individual patients are likely to benefit from injections or whether different treatment options should be prioritised. Clearly, it can be debated what level of accuracy is needed to support treatment selection. Our analyses indicate that pain intensity 2 weeks after the injection can be predicted with a median error of 1.36 on a 0 to 10 NRS. While this level of accuracy may already allow for triaging patients, more accurate predictions might be achieved in the future by augmenting the prediction models with further information (eg, clinical history or neurophysiological data). In this study, we deliberately focused on simple models that only required a minimum of information.

A few additional results are worth highlighting. First, adding information about individual beliefs obtained directly after the injection (T_2) improved neither explanation nor prediction of outcomes. We had expected this postintervention information to be potentially valuable because the immediate effect of the injection (due to the local anesthetic) might induce an additional expectation that could influence longer-term outcome. However, our results speak against this. Furthermore, in separate analyses, we examined which factors might contribute to the formation of beliefs about treatment outcome. Among the 5 factors, we considered—self-efficacy, previous injections and their perceived success, pain duration, weight of the opinions of others in the decision to undergo injection therapy, and educational background—the opinions of others showed a significant relationship with belief certainty. Surprisingly, however, none of the above factors showed a significant association with expectations about treatment outcome. Finally, we performed a sensitivity analysis in order to examine whether our results depended on the specific questionnaire we used. To this end, we repeated our analyses using scores from the ETS questionnaire. Reassuringly, both the within-sample analyses of associations and the out-of-sample prediction analyses provided the same qualitative results.

Our study has strengths and weaknesses. Beginning with strengths, we focused on a well-defined intervention in a single-center setting that provided a relatively homogenous clinical situation and a precise time point at which expectations can be assessed. Our sample size was informed by a conservative power analysis, and we obtained multiple measurements of key components of individual beliefs (expectations and certainty), as a basis for sensitivity analyses.

Turning to weaknesses, we used 2 measures of belief precision (VAS and a questionnaire asking to exclude possible future pain states) that have not been validated. Generally, the question of how to measure subjective belief precision or certainty is an important topic in cognitive neuroscience and is beginning to be studied in pain research.[10] It is worth mentioning that the construct of "confidence" is often used interchangeably; here, we follow previous recommendations[36] and refer to "certainty" as opposed to "confidence." For example, in studies of metacognition, indices of certainty are often computed using signal detection theory or Bayesian models, based on trial-wise responses (eg, using VAS) about the certainty in one's decisions.[13,14] To our knowledge, the construct and predictive validity of scales measuring subjective certainty are not established. This represents an important area of future work.

A second important limitation of our work is the lack of a completely separate (held-out) test data set. External validation of our current approach, using data sets from additional samples, will be important.

Finally, our current study was restricted to data from simple and brief questionnaires. While a deliberate choice for this study, including additional biological data might substantially increase the accuracy of outcome predictions.[42] In particular, neuroimaging and electrophysiological data have proven powerful for predictions in other areas of pain research.[3,28,29,41,44] Clearly, their additional complexity and costs make clinical applications considerably less straightforward.

In summary, this study demonstrates that treatment expectations are not only associated with 2-week outcomes after lumbar injection therapy but also allow for individual predictions. Future studies should examine whether such predictions could usefully guide treatment selection. Additionally, our findings suggest that injection therapy could be enhanced by cognitive interventions targeting expectations. Finally, it would be worthwhile investigating the possibility of including neurophysiological or neuroimaging data into predictive models of treatment outcomes, while carefully considering practical feasibility and the cost-benefit ratio.

Conflict of interest statement

The authors have no conflicts of interest to declare.

Acknowledgements

This work was supported by the Clinical Research Priority Program "Pain" at the University of Zurich (C.W. and K.E.S.), the Wilhelm Schulthess Foundation (Z.M.M.), and the René and Susanne Braginsky Foundation (K.E.S.).
K.W. has received consultancy fees from P&G Health, Germany. The remaining authors have no conflicts of interest to declare.
The analyses tested in this paper were specified ex ante in a time-stamped analysis plan (https://gitlab.ethz.ch/tnu/analysis-plans/muellerschraderetal_pain_expectation_2021). The analysis code is publicly available on (https://gitlab.ethz.ch/tnu/code/muellers-chraderetal_pain_expectation_2021). All patients were asked for consent that their data could be shared; for those patients who agreed, the data which entered the analyses presented in this manuscript can be obtained on request from tnu-datasharing@biomed.ee.ethz.ch after signing a data use agreement.

Appendix A. Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PAIN/B636.

Article history:
Received 22 October 2021
Received in revised form 1 April 2022

Accepted 15 April 2022
Available online 10 May 2022

References

[1] Anchisi D, Zanon M. A Bayesian perspective on sensory and cognitive integration in pain perception and placebo analgesia. PLoS One 2015;10: e0117270.
[2] Auer CJ, Glombiewski JA, Doering BK, Winkler A, Laferton JAC, Broadbent E, Rief W. Patients' expectations predict surgery outcomes: a meta-analysis. Int J Behav Med 2016;23:49–62.
[3] Baliki MN, Petre B, Torbey S, Herrmann KM, Huang L, Schnitzer TJ, Fields HL, Apkarian AV. Corticostriatal functional connectivity predicts transition to chronic back pain. Nat Neurosci 2012;15:1117–19.
[4] Barth J, Kern A, Lüthi S, Witt CM. Assessment of patients' expectations: development and validation of the Expectation for Treatment Scale (ETS). BMJ Open 2019;9:e026712.
[5] Bingel U. Placebo 2.0: the impact of expectations on analgesic treatment outcome. PAIN 2020;161:S48–56.
[6] Bingel U, Wanigasekera V, Wiech K, Ni Mhuircheartaigh R, Lee MC, Ploner M, Tracey I. The effect of treatment expectation on drug efficacy: imaging the analgesic benefit of the opioid remifentanil. Sci Translational Med 2011;3:70ra14.
[7] Brown TT, Lee W. The FUTUREPAIN study: validating a questionnaire to predict the probability of having chronic pain 7-10 years into the future. PLoS One 2020;15:e0237508.
[8] Büchel C, Geuter S, Sprenger C, Eippert F. Placebo analgesia: a predictive coding perspective. Neuron 2014;81:1223–39.
[9] Cormier S, Lavigne GL, Choinière M, Rainville P. Expectations predict chronic pain treatment outcomes. PAIN 2016;157:329–38.
[10] Dildine TC, Necka EA, Atlas LY. Confidence in subjective pain is predicted by reaction time during decision making. Scientific Rep 2020;10:21373.
[11] Doering BK, Glombiewski JA, Rief W. Expectation-focused psychotherapy to improve clinical outcomes. Int Rev Neurobiol 2018; 138:257–70.
[12] Fadyl J, McPherson K. Return to work after injury: a review of evidence regarding expectations and injury perceptions, and their influence on outcome. J Occup Rehabil 2008;18:362–74.
[13] Fleming SM. HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. Neurosci Conscious 2017;2017: nix007.
[14] Fleming SM, Lau HC. How to measure metacognition. Front Hum Neurosci 2014;8:443.
[15] Friston K. A theory of cortical responses. Philosophical Trans R Soc B 2005;360:815–36.
[16] Geuter S, Boll S, Eippert F, Büchel C. Functional dissociation of stimulus intensity encoding and predictive coding of pain in the insula. eLife 2017; 6:e24770.
[17] Grahl A, Onat S, Büchel C. The periaqueductal gray and Bayesian integration in placebo analgesia. eLife 2018;7.
[18] Hallegraeff JM, Krijnen WP, van der Schans CP, de Greef MH. Expectations about recovery from acute non-specific low back pain predict absence from usual work due to chronic low back pain: a systematic review. J Physiother 2012;58:165–72.
[19] Hayden JA, Wilson MN, Riley RD, Iles R, Pincus T, Ogilvie R. Individual recovery expectations and prognosis of outcomes in non-specific low back pain: prognostic factor review. Cochrane database Syst Rev 2019; 2019.
[20] Hoskin R, Berzuini C, Acosta-Kane D, El-Deredy W, Guo H, Talmi D. Sensitivity to pain expectations: a Bayesian model of individual differences. Cognition 2019;182:127–39.
[21] Iles RA, Davidson M, Taylor NF, O'Halloran P. Systematic review of the ability of recovery expectations to predict outcomes in non-chronic non-specific low back pain. J Occup Rehabil 2009;19:25–40.
[22] Jung WM, Lee YS, Wallraven C, Chae Y. Bayesian prediction of placebo analgesia in an instrumental learning model. PLoS One 2017;12:e0172609.
[23] Kampermann L, Tinnermann A, Büchel C. Generalization of placebo pain relief. PAIN 2021;162:1781–9.
[24] Keller GB, Mrsic-Flogel TD. Predictive processing: a canonical cortical computation. Neuron 2018;100:424–35.
[25] Kersten D, Mamassian P, Yuille A. Object perception as Bayesian inference. Annu Rev Psychol 2004;55:271–304.
[26] Kongsted A, Vach W, Axø M, Bech RN, Hestbaek L. Expectation of recovery from low back pain: a longitudinal cohort study investigating patient characteristics related to expectations and the association between expectations and 3-month outcome. Spine 2014;39:81–90.
[27] Kube T, Glombiewski JA, Rief W. Using different expectation mechanisms to optimize treatment of patients with medical conditions: a systematic review. Psychosomatic Med 2018;80:535–43.
[28] Kutch JJ, Labus JS, Harris RE, Martucci KT, Farmer MA, Fenske S, Fling C, Ichesco E, Peltier S, Petre B, Guo W, Hou X, Stephens AJ, Mullins C, Clauw DJ, Mackey SC, Apkarian AV, Landis JR, Mayer EA. Resting-state functional connectivity predicts longitudinal pain symptom change in urologic chronic pelvic pain syndrome: a MAPP network study. PAIN 2017;158:1069–82.
[29] Lee JJ, Kim HJ, Čeko M, Park BY, Lee SA, Park H, Roy M, Kim SG, Wager TD, Woo CW. A neuroimaging biomarker for sustained experimental and clinical pain. Nat Med 2021;27:174–82.
[30] Mangels M, Schwarz S, Sohr G, Holme M, Rief W. Der Fragebogen zur Erfassung der schmerzspezifischen Selbstwirksamkeit (FESS). Diagnostica 2009;55:84–93.
[31] Ongaro G, Kaptchuk TJ. Symptom perception, placebo effects, and the Bayesian brain. PAIN 2019;160:1–4.
[32] Peerdeman KJ, van Laarhoven AI, Peters ML, Evers AW. An integrative review of the influence of expectancies on pain. Front Psychol 2016;7:1270.
[33] Peerdeman KJ, van Laarhoven AIM, Keij SM, Vase L, Rovers MM, Peters ML, Evers AWM. Relieving patients' pain with expectation interventions: a meta-analysis. PAIN 2016;157:1179–91.
[34] Petzschner FH, Glasauer S, Stephan KE. A Bayesian perspective on magnitude estimation. Trends Cognitive Sci 2015;19:285–93.
[35] Petzschner FH, Weber LAE, Gard T, Stephan KE. Computational psychosomatics and computational psychiatry: toward a joint framework for differential diagnosis. Biol Psychiatry 2017;82:421–30.
[36] Pouget A, Drugowitsch J, Kepecs A. Confidence and certainty: distinct probabilistic quantities for different goals. Nat Neurosci 2016;19:366–74.
[37] R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
[38] Schafer SM, Geuter S, Wager TD. Mechanisms of placebo analgesia: a dual-process model informed by insights from cross-species comparisons. Prog Neurobiol 2018;160:101–22.
[39] Stephan KE, Manjaly ZM, Mathys CD, Weber LA, Paliwal S, Gard T, Tittgemeyer M, Fleming SM, Haker H, Seth AK, Petzschner FH. Allostatic self-efficacy: a metacognitive theory of dyshomeostasis-induced fatigue and depression. Front Hum Neurosci 2016;10:550.
[40] Tabor A, Thacker MA, Moseley GL, Körding KP. Pain: a statistical account. PLoS Comput Biol 2017;13:e1005142.
[41] Tétreault P, Mansour A, Vachon-Presseau E, Schnitzer TJ, Apkarian AV, Baliki MN. Brain connectivity predicts placebo response across chronic pain clinical trials. PLOS Biol 2016;14:e1002570.
[42] van der Miesen MM, Lindquist MA, Wager TD. Neuroimaging-based biomarkers for pain: state of the field and current directions. Pain Rep 2019;4:e751.
[43] Wager TD, Atlas LY. The neuroscience of placebo effects: connecting context, learning and health. Nat Rev Neurosci 2015;16:403–18.
[44] Wager TD, Atlas LY, Lindquist MA, Roy M, Woo CW, Kross E. An fMRI-based neurologic signature of physical pain. New Engl J Med 2013;368:1388–97.
[45] Wiech K. Deconstructing the sensation of pain: the influence of cognitive processes on pain perception. Science 2016;354:584–7.