



OPEN

# Prognostic subgroups of chronic pain patients using latent variable mixture modeling within a supervised machine learning framework

Xiang Zhao<sup>1</sup>, Katharina Dannenberg<sup>2</sup>, Dirk Repsilber<sup>2</sup>, Björn Gerdle<sup>3,5</sup>, Peter Molander<sup>3,4</sup> & Hugo Hesser<sup>1,4</sup>✉

The present study combined a supervised machine learning framework with an unsupervised method, finite mixture modeling, to identify prognostically meaningful subgroups of diverse chronic pain patients undergoing interdisciplinary treatment. Questionnaire data collected at pre-treatment and 1-year follow up from 11,995 patients from the Swedish Quality Registry for Pain Rehabilitation were used. Indicators measuring pain characteristics, psychological aspects, and social functioning and general health status were used to form subgroups, and pain interference at follow-up was used for the selection and the performance evaluation of models. A nested cross-validation procedure was used for determining the number of classes (inner cross-validation) and the prediction accuracy of the selected model among unseen cases (outer cross-validation). A four-class solution was identified as the optimal model. Identified subgroups were separable on indicators, predictive of long-term outcomes, and related to background characteristics. Results are discussed in relation to previous clustering attempts of patients with diverse chronic pain conditions. Our analytical approach, as the first to combine mixture modeling with supervised, targeted learning, provides a promising framework that can be further extended and optimized for improving accurate prognosis in pain treatment and identifying clinically meaningful subgroups among chronic pain patients.

**Keywords** Pain classification, Latent variable mixture modeling, Machine learning, Pain prognosis

Subgrouping of patients with chronic pain has been an enduring research topic as it provides information for clinical assessment, accurate prognosis, and personalized healthcare<sup>1,2</sup>. The negative outcomes of chronic pain—pain that has lasted over three months—include disability, poor quality of life, and high health care consumption<sup>3,4</sup>. Clinical characteristics, presentation, and functioning among patients suggest significant individual heterogeneities<sup>2,5–8</sup>. Given that multiple biopsychosocial factors are known to contribute to chronic pain, interdisciplinary treatment is commonly the treatment of choice for most patients<sup>4,9</sup>. Yet, while findings indicate that patients with the severest clinical presentation seem to benefit most from such a treatment<sup>10</sup>, it is still unclear which patient- and treatment-related factors contribute to the treatment success<sup>11</sup>.

Data-driven approaches use large-scale datasets and rich information for decision making, providing a unique opportunity to identify patterns of information that can be used to improve classification and prediction in clinical health research<sup>12</sup>. Given the notable heterogeneity in both clinical presentation and treatment response, several studies have also been devoted to identifying clinically relevant subgroups using data-driven approaches in specific pain diagnoses<sup>13,14</sup>.

Although it is increasingly recognized that patients with a certain chronic pain diagnosis are heterogeneous in clinical presentation<sup>8</sup>, less attention has been paid to identifying subgroups among a broader population of

<sup>1</sup>School of Behavioural, Social and Legal Sciences, Örebro University, Fakultetsgatan 1, 702 81 Örebro, Sweden. <sup>2</sup>School of Medical Sciences, Örebro University, Örebro, Sweden. <sup>3</sup>Department of Health, Medicine and Caring Sciences, Pain and Rehabilitation Centre, Linköping University, Linköping, Sweden. <sup>4</sup>Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden. <sup>5</sup>Björn Gerdle is deceased. ✉email: Hugo.Hesser@oru.se

chronic pain patients. Yet, a handful of studies have examined subgroups in patients receiving interdisciplinary treatment, providing initial evidence for clinically relevant subgroups across diverse pain conditions<sup>10,15–17</sup>.

There are two major drawbacks of the current pain classification research. First, psychological and social dimensions are not well incorporated into the classification processes, even though a substantial body of evidence suggests that psychological and social aspects are important for the development, maintenance and treatment of pain<sup>4,18,19</sup>. Clustering attempts (i.e., identifying subgroups) have either been focused on specific psychological or biomedical aspects of pain, with only few studies having combined psychosocial dimensions and pain characteristics (e.g., pain extent, pain intensity)<sup>20</sup>. Second, the analytical approach to identifying subgroups has a large space to advance. Specifically, early subgroup profiling attempts used a limited number of indicator variables to form subgroups or primitive subgrouping methods (e.g., by using arbitrary cut-off points on a single measure or by using simple regression models). Thus, better categorization attempts are needed especially to improve the prediction of response to treatment and the matching of patients to treatments<sup>21</sup>.

In contrast to these early subgroup identifying approaches, advanced data-driven analytical methods aiming to uncover subgroups in heterogeneous populations may hold more promise. Latent variable mixture modeling (LVMM) is an advanced person-centered analytic method, which enables researchers to classify individuals into unobserved (latent) subpopulations based on patterns of individual characteristics<sup>22,23</sup> and may offer valuable information for individual level empirical classification and prediction<sup>16,24</sup>. Although underutilized in pain research overall the method has more recently gained increased attention in the field<sup>24</sup>.

While LVMM is a promising method, there are also some key analytical challenges when using LVMM, like any other data-driven approach, for the purpose of identifying clinically meaningful subgroups in a heterogeneous population, such as chronic pain patients. In the applied sciences, LVMM is frequently used in an exploratory manner, with model selection relying primarily on comparing several models in terms of their quality of fit. Overfitting is the primary concern with any exploratory model, and this is also true for LVMM. Furthermore, no single criterion can be used to determine optimal fit, so researchers will have to rely on a combination of statistical and pragmatic criteria to select a final model. In fact, derived subgroups from LVMM have been difficult to replicate in independent samples and, given the considerable uncertainty over model selection, a general recommendation is not to treat the “best-fitted model” for the sample as the only model for the population<sup>23</sup>. Furthermore, methodologists have emphasized the necessity of “substantive checking”, such as linking derived classes to auxiliary information, in determining whether the best-fitted class-solution is empirically and theoretically meaningful<sup>25</sup>. Yet, while auxiliary variables (e.g., predictors or distal outcomes) can readily be incorporated in LVMM to validate the identified subgroups, evaluation and selection of the model is generally not determined by any external performance (e.g., prediction accuracy). Hence, subgroups identified through this approach may not provide useful prognostic information.

Both aforementioned concerns regarding the use of LVMM could potentially be combated within a machine learning framework. First, machine learning researchers have long acknowledged the susceptibility of highly flexible algorithms to overfitting the sample data; consequently, numerous approaches to both assessing and preventing overfitting (e.g., cross-validation) have been developed. Second, supervised learning methods incorporate an external criterion by which the classification model should be evaluated and optimized. In fact, LVMM has recently been formally integrated into a supervised learning framework with the main goal of evaluating candidate LVMM by linking them to clinically relevant auxiliary information<sup>26</sup>. In other words, rather than searching for the “true” model based on internal fit criteria, the primary purpose is to select a model based on its predictability. Indeed, accurately predicting or classifying individual-level outcomes is critical for enhancing the quality of personalized treatment and intervention. Thus, approaching subgrouping and pain classification from this prediction perspective may be preferable.

The main aim of the present study was to identify prognostically meaningful subgroups of chronic pain patients. Hence, the present study used a supervised machine learning framework (i.e., cross validation) for model selection and performance evaluation of LVMM. Specifically, with national real-world clinic data from chronic pain patients in Sweden, we identified models that capture individual heterogeneity in phenotypic characteristics using LVMM. Patients were grouped based on scores on several measures of pain characteristics, psychological aspects, and social functioning and general health status. For both model selection and performance evaluation, we used the self-reported pain interference one-year following the interdisciplinary treatment as the outcome; thus, by using this external criterion we could evaluate the empirically derived subgroups based on a direct measure of success (i.e., the predictability of the future outcome). To evaluate the expected accuracy of the selected model (i.e., class solution) on new data (i.e., unseen cases), we implemented a systematic nested cross-validation approach.

## Methods

### Participants and procedure

A sample of patients with complex chronic nonmalignant pain who were referred to specialist care centers from the Swedish Quality Registry for Pain Rehabilitation (SQRP) was used in this study. With Swedish national data, the SQRP provides important information from patients whose chronic pain is not caused by any other condition in need of urgent treatment<sup>15</sup>. The treatment received varies across clinical and patients in terms of both length and content: however, the treatment is typically delivered in a group format by an interdisciplinary team over several weeks to a few months, with a focus on pain education, supervised physical training and psychological interventions based on principles from cognitive behavioral therapy. Detailed description of the registry and collected data are presented elsewhere<sup>15</sup>.

During the data cleaning stage, we kept the cases with valid answers on follow-up pain intensity and interference, as well as removed 349 cases with duplicated identification numbers (potentially due to repeated

registrations), resulting in a final subsample with 11,995 patients ( $M_{\text{age}} = 43.9$  years; 77.6% women; 46% full-time employed; 82% Swedish born). The patients in this sample undertook their initial visits to a specialist unit for their pain issues from 2008 to 2016.

The baseline data were collected at the specialist unit on the patients' initial visits. The follow-up data were subsequently accessed one year after the rehabilitation. All subjects and/or their legal guardian(s) gave their written and informed consent during their first visit. This study was approved by the Swedish Ethical Review Authority, 2018-036, EPM dnr: 2019-02167, EPM 2020-02038. It was performed in accordance with the Declaration of Helsinki.

## Measures

We selected our indicator variables mainly based on mainstream recommendations for pain treatments<sup>2,27,28</sup> as well as findings from previous studies using SQRP for classification and prediction. The selected indicator variables comprise three domains: (1) pain characteristics, (2) psychological aspects, and (3) social functioning and general health status. The survey was undertaken in Swedish.

### Indicator variables

**Pain characteristics.** Two widely used pain aspects from the Multidimensional Pain Inventory<sup>29</sup> were used: *pain intensity* (2-item) refers to the magnitude of experienced pain, *pain interference* (11-item) reflects interference of pain in one's work and social functioning. This inventory has shown good psychometric properties in previous Swedish studies<sup>30</sup>. The average score of each subscale represents the construct. With a 7-point scale (from *never* [0] to *very often* [6]), higher scores reflect stronger pain intensity, interference, and more life control, respectively. In addition, we also used the single-item *pain location number* (i.e., the number of anatomical regions with pain) as it appears to be a key correlate of pain aspects and indicates the pain spreading extent<sup>31</sup>. To quantify pain location numbers, a list of 36 predefined anatomical areas were presented to participants who then indicated their painful areas<sup>31</sup>. The number of pain locations ranged from 0 to 36, larger numbers indicate greater pain extent.

**Psychological aspects.** *Anxiety* and *depression* were included as previous studies suggest that these two variables are related to pain conditions as well as analgesic treatment outcomes<sup>2</sup>. We assessed these two constructs using the 14-item Hospital Anxiety and Depression Scale<sup>32</sup>, with a 4-point scale (e.g., from *not at all* [0] to *very often* [3]). The Swedish version<sup>33</sup> has been found as a reliable instrument. Higher scores suggest stronger anxiety and depressive symptoms. The *affective distress* (3-item) from the Multidimensional Pain Inventory<sup>29</sup> was also included as an indicator variable as it captures more general distress symptoms. Informed by the fear-avoidance model<sup>34</sup>, we further included *fear of movement*, which is considered as an anticipatory emotional response to a future bodily movement that could lead to painful feeling. Fear of movement was measured using the 17-item Tampa Scale for Kinesiophobia<sup>35</sup>, where each item is rated on a 4-point scale (from *strongly disagree* [1] to *strongly agree* [4]). A previous study<sup>36</sup> showed the good reliability of this instrument among Swedish pain patients. Item scores were added up to form the total construct score; higher scores reflect more perceived fear of physical movements. In addition, *mental health* (5-item), a subscale from the Short Form Health Survey (SF-36)<sup>37,38</sup> was included to reflect the overall mental health status.

**Social functioning and general health status.** Health-related quality of life has been suggested as another core criterion to assess in pain treatment<sup>28</sup>. Based on the Short Form Health Survey (SF-36)<sup>37,38</sup>, we specifically chose the overall *physical health* (21-item), which includes general health, physical functioning, role limitations due to physical problems, and body pain. The subscale *vitality* (4-item) was also included. While mental wellbeing has been suggested as a key assessment measure<sup>28</sup>, vitality has not been widely recommended as a key factor for analgesic outcome until recently<sup>2</sup>. Vitality plays an important role among pain patients as physical and mental energy is required to manage pain<sup>39</sup>. Items of each subscale were transformed and summed up as per the scale instructions<sup>40</sup>. Higher scores reflect more perceived vitality and physically healthier status, respectively. In addition, *life control* (4-item), which indicates how much control one generally has over their life, was selected from the Multidimensional Pain Inventory<sup>29</sup>.

### Outcome variable

Two variables assessed at the one-year follow-up stage were used as distal outcome variables: (1) *pain interference* (11-item), a subscale from the Multidimensional Pain Inventory<sup>29</sup> and (2) *health status*, a single-item measure from the EQ-5D<sup>41</sup>. The pain interference scale properties are described above. For the health status measure, participants were asked to rate their general health status on a vertical analogue scale from 0 to 100, with 100 indicating the best possible health level. Pain interference was used in the cross-validation process because pain interference reflects overall impact of pain and serves as a primary concern of patients and represents a core outcome domain<sup>42,43</sup>.

### Background variables

To further understand the characteristics of each phenotype, we included sex, age, highest education level, employment status (full-time/part-time), visits to a physician about pain in the past year (0-1/2-3/over 4 times), pain duration (i.e., years of pain experience), and social support. Demographic and patient disposition factors were included as per previous recommendations<sup>27,28</sup>. As social support has been found as a key variable for pain recovery<sup>18,20</sup>, the social support subscale (2-item) from the Multidimensional Pain Inventory<sup>29</sup> was used.

## Statistical analysis

### *Finite mixture model*

The analytical model that was evaluated was a Gaussian finite mixture model. The model uses the information from the observed indicator variables to form subgroups where the grouping variable is formalized as a latent categorical variable with latent classes from  $c_1$  to  $c_k$ . A mixture distribution is a weighted sum of  $K$  component distributions as follows,

$$f(y_i) = \sum_{k=1}^K \pi_k f_k(y_i, \theta_k)$$

where  $y_i$  is a vector of observed indicator variables for individual  $i$ ,  $\pi_k$  is a weight that quantifies the  $k^{\text{th}}$  component's relative size, and  $\theta_k$  is a vector of model parameters for the  $k^{\text{th}}$  component<sup>44</sup>. The component distributions,  $f_k$ , are the multivariate normal distributions, where the parameter vector  $\theta_k$  contains the parameters with the class-specific means,  $\mu_k$ , and covariance matrices,  $\Sigma_k$ . The off-diagonal elements of  $\Sigma_k$  matrices were fixed to zero (i.e., local independence assumption), and diagonal elements were constrained to be equal across classes. One can think of  $f_k$  as the conditional probability of obtaining the observed sample data/indicator values for an individual from class  $k$ . Then,  $f(y_i)$  gives the unconditional probability for  $y_i$ . We fitted a successively increasing number of classes,  $k = \{2, 3, 4, 5, 6\}$ . Models were evaluated with Mplus using maximum likelihood estimation and the EM-algorithm<sup>45</sup>. To avoid potential convergence at local maxima, we used random starting values (250 for the initial and 100 for the final stage optimization in Mplus). Posterior probabilities for each individual and class were computed based on model parameters ( $\mu_k$ ,  $\Sigma_k$ ) and the observed individual score of each indicator variable [see equations 32 and 33<sup>46</sup>]. These computed probabilities were then used as predictors ( $k - 1$ ) of the pain interference outcome at one-year follow-up, based on a linear regression model in the nested cross-validation procedure.

Although there are no formal tests for determining the assumption of normally distributed variables within each latent class in Gaussian finite mixture models, we visually examined the data distributions of each indicator for the whole sample and each class in the final class solution, finding that normality can be generally assumed across most distributions, although pain intensity, pain interference, and vitality seemed to have larger skewness.

### *Combined model for outcome prediction*

Posterior class probabilities were computed for each individual in test sets of our cross validation and used for outcome prediction based on a linear regression model. The parameters for this regression model are obtained in the respective training set. This final prediction step constitutes the supervised part of our machine learning algorithm and hence enables constructing latent classes with the perspective of predicting a distal outcome.

### *Cross-validation for parameter optimization and performance assessment*

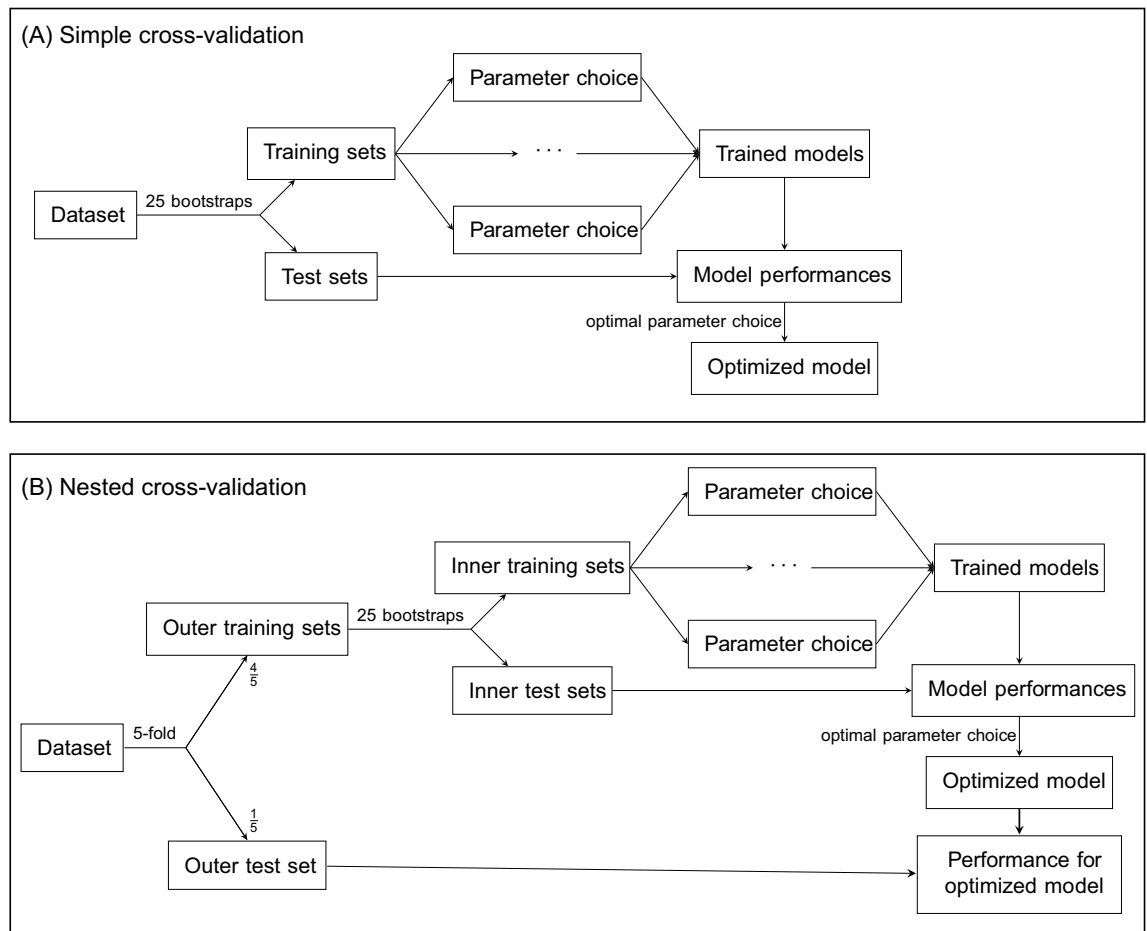
A cross-validation is performed whenever a model is fitted, to determine a suitable value for the parameters of the model, in our case the number  $k$  of subgroups (or the number of subgroups,  $k$ ), based on prediction accuracy<sup>47–49</sup>. As detailed in Fig. 1A, a bootstrap approach is employed to repeatedly resample a training set from the data (predictors and outcome). Data not chosen during the bootstrap form the test set. Different parameter settings are applied to train a model for each round of the cross-validation, and each time the trained model's performance is assessed using the test data. Model performance on these test set data was quantified as root mean squared error (RMSE). Finally, from average model performances for different parameter choices, the optimal parameter choice can be concluded. A final model is then fit to the entire dataset using this optimal parameter choice (i.e., yielding the optimized model).

Using a simple cross-validation, as described here above, a prediction model can be optimized, but its performance cannot be assessed at the same time. Therefore, we employed a nested cross-validation<sup>50–52</sup>, as shown in Fig. 1B: The available data are split into 5 parts, and 5 models are trained and optimized, each one leaving out one fifth of the data for measuring the performance. The other four fifths of the data (the outer training set) are split up further in the inner cross-validation (as in Fig. 1A). The performance of each optimized model can then be assessed using the share of the data that was left out in the training of the model, such that no sample is ever used to learn and evaluate the same model. Model performance on these outer test set data was again quantified as RMSE. Importantly, as it is not possible to fit any model to all available data, and, at the same time, estimate its performance on future new data, the average performance assessed in the five outer test-set predictions of our combined LVMM-linear models served as our best estimate for the future performance of such a “final” model. To be able to compare the results of this selection process to a conventional procedure for model selection we provide commonly used fit statistics used in mixture modeling in supplementary materials (Table S3).

Specifically, we applied a five-fold outer cross-validation in combination with a 25-fold bootstrap for the inner cross-validation. The R-package “MplusAutomation” (version 1.1.0)<sup>53</sup> was used as an interface from R to Mplus (version 8.5)<sup>54</sup>. Our approach was then implemented as a custom model using the R-package “caret” (version 6.0.91)<sup>55</sup>, which provides the functionality for the inner cross-validation.

### *Subgroup characteristics and associations with auxiliary information*

Based on the performance of models, the optimal latent class solution was then estimated using the whole dataset. The consequent subgroup profiling and comparisons were based on the optimal model (i.e., optimal  $k$ ) selected from the cross-validation (as described above). Estimated means harvested from the whole sample-based model were plotted to profile the characteristics of each subgroup. To investigate how subgroup memberships related to both background variables and distal outcome (one-year follow-up pain interference), the three-step approaches were used in order not to allow the auxiliary information to influence the class-solution at this descriptive stage<sup>56</sup>. Effect sizes (Cohen's  $d$ ) for outcomes were computed when applicable.



**Figure 1.** Workflow for two cross-validation schemes.

Following the assumption of missing at random, the full information maximum likelihood approach was adopted for most analyses; three cases were removed since they had missing values on all indicators. Listwise deletion was used when examining the relationships between background covariates and class membership, with 9173 cases (76.5%).

## Results

### Cross-validation for parameter optimization and performance assessment

#### Model selection

Following the five-fold nested cross-validation, the optimal number of latent classes was chosen based on two- to six-class models in the inner cross-validation and their performances measured on the outer test sets. The resulting RMSE values of each model with respect to simplicity (of the model) and prediction of the outcome suggested the four-class model as the optimal (Fig. 2). Notably, there is one large relevant drop in RMSE values, from the three-class to the four-class model ( $\Delta_{\text{RMSE}} = 0.072$ ). Similar trends were observed in other performance indices including  $R^2$  and mean square error; for brevity, these values are tabulated in supplementary materials (Table S1).

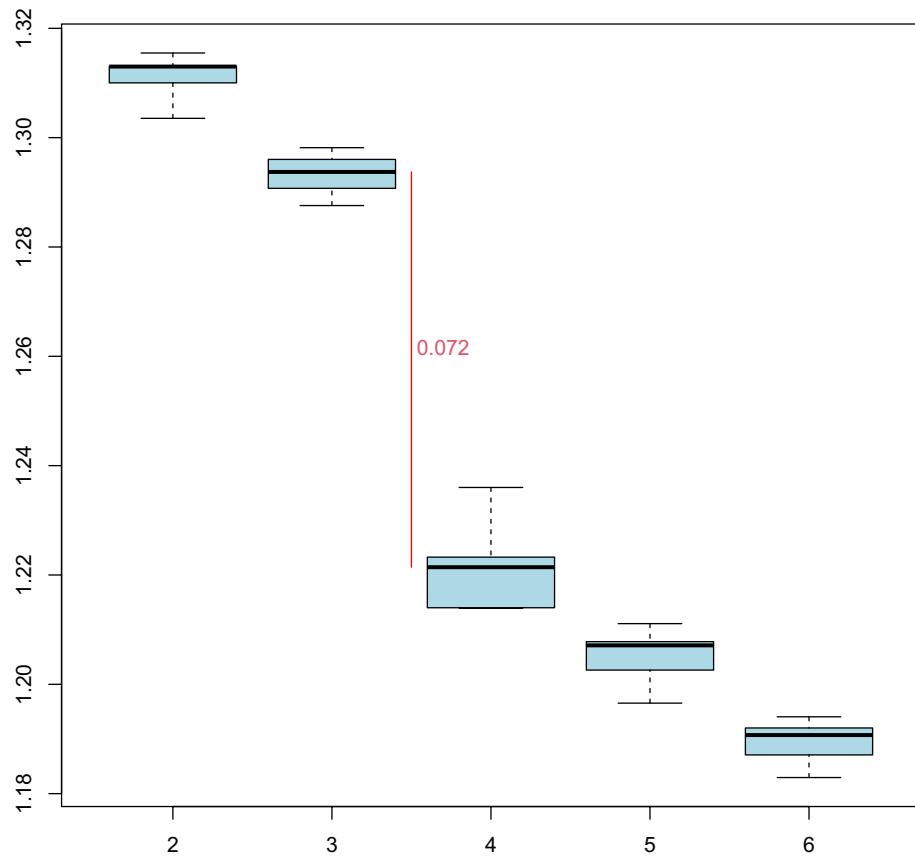
#### Model performance

The performance of the selected model ( $k = 4$ ) was evaluated in five outer cross-validation rounds as five-fold cross-validation, such that each sample belonged to one of the five outer test sets and was predicted once by a model which was trained on 80% of the total data. As assessed with a linear regression, the predicted pain interference outcome value by the latent classes (i.e., posterior probabilities) captured about 26% of the variation in the true pain interference values of the dataset. Of note, performance of the combined LVMM-linear model was evaluated by (outer) cross-validation, namely, using the fitted model to predict the outcome for unseen data (i.e., the five outer test sets comprising 20% of all data each which have not been used for model selection purpose).

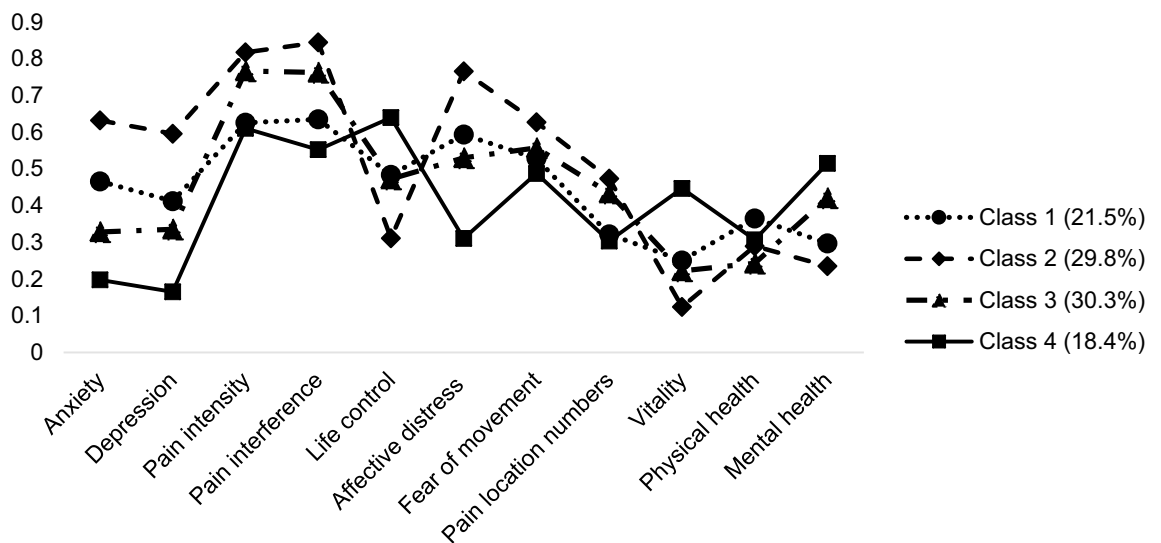
### Optimized final mixture model fitted to all data

#### Profile of subgroups

Informed by the previous steps, an LVMM with  $k = 4$  was performed in the whole sample, resulting in four qualitatively distinct subgroups. Figure 3 visually illustrates the characteristics of subgroups (for tabulated information, see Table S2 in supplementary materials).



**Figure 2.** The Root-Mean-Square deviation (RMSE) deflection with four-class solution. *Note* The boxplots of the RMSE values (y-axis) for the five cross validations, for different subgroup numbers (x-axis). For each number of subgroups, the boxplots show the minimum, 1st, 2nd (median), 3rd quartiles, and the maximum of the RMSE values of the outer cross validations and their medians as horizontal black lines. The red vertical line denotes the maximum drop of the RMSE with reflected median difference between subgroup number 3 and 4.



**Figure 3.** Estimated means of the four-class model generated from latent variable mixture modeling. *Note* To visually compare between indicators with various ranges, estimated means ( $\bar{x}$ ) of each indicator have been rescaled based on the theoretical maximum value ( $max_x$ ) of each scale (i.e.,  $\frac{\bar{x}}{max_x}$ ). A larger rescaled score reflects a higher rank of an estimated mean.



Class 1 comprised 2573 patients (21.5%) who reported relatively low pain intensity, interference, and pain location numbers. Although Class 1 members showed the best level of general physical health, their negative emotions such as anxiety and depression were notably higher.

With 3575 patients (29.8%), Class 2 demonstrated the worst levels on most indicators. Specifically, members of Class 2 reported severest pain intensity, interference, as well as pain extent. Their anxiety, depression, and general affective distress were also highest among the subgroups. Moreover, these patients had markedly lower mental health and vitality scores.

Class 3 was the largest subgroup (30.3%), with 3638 patients. Patients in Class 3 had the lowest level of physical health, although their negative emotions were relatively moderate compared with other subgroups. Even though Class 3 patients showed a similarly high level in pain intensity, interference, and pain extent as in Class 2, they had a relatively more life control than Class 2 members.

Class 4 ( $n = 2206$ ; 18.4%) showed the best scores on most indicators. Their pain presentation and negative emotions were the lowest, with a better health condition especially mental health and vitality. Based on these subgroup features, Class 4 was used as the reference group in the subsequent analyses as it is regarded as the healthiest subgroup.

#### Subgroup associations with follow-up outcomes

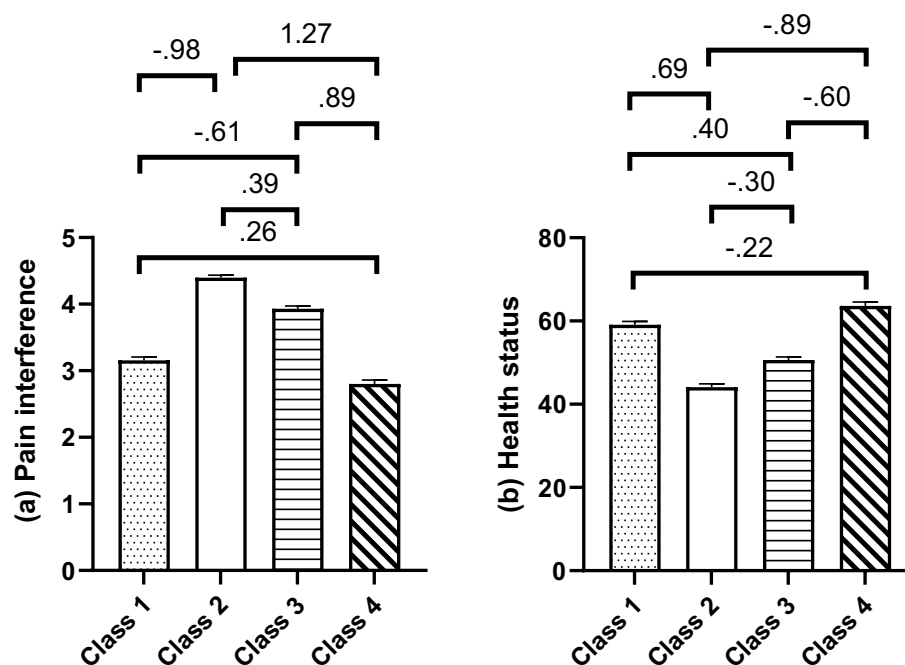
With the three-step approach which included distal outcomes, the between-subgroup difference in the follow-up pain interference ( $\chi^2_{(3)} = 3545.30$ ,  $p < 0.001$ ) and health status ( $\chi^2_{(3)} = 1172.73$ ,  $p < 0.001$ ) both showed an overall statistical significance. Figure 4 shows the means for each subgroup and the standardized mean differences (Cohen's  $d$ ) for all group comparisons.

For pain interference, all between-subgroup differences were statistically significant ( $ps < 0.001$ ) except for the comparison between Classes 1 and 4 ( $p = 0.130$ ). Large effect sizes ( $d \geq 0.80$ ) were found in Class 2 vs. Class 4 ( $d = 1.27$ ), Class 1 vs. Class 2 ( $d = -0.98$ ), and Class 3 vs. Class 4 ( $d = 0.89$ ).

For health status, all between-subgroup differences showed a statistical significance ( $ps < 0.001$ ; Class 1 vs. Class 4,  $p = 0.001$ ). Large effect size ( $d \geq 0.80$ ) was found in Class 2 vs. Class 4 ( $d = -0.89$ ).

#### Subgroup associations with background characteristics

Table 1 descriptively reports the background variables of the four subgroups based on the most likely class membership of each subject (i.e., largest posterior probabilities). When background variables were included as auxiliary predictors, all variables showed associations with the class membership (Table 2). Compared with the healthiest subgroup (Class 4), other classes were younger. Overall, all subgroups had more women than men (i.e., > 70% women); between-subgroup differences were found in sex distributions. Patients in Classes 2 and 3 had more women and were less educated than Class 4, whereas those in Class 1 were more educated than Class 4. Furthermore, as opposed to Class 4, members in Classes 2 and 3 were less full-time employed, had more visits to a pain clinic in the past year, as well as had longer pain durations. While patients in Class 1 showed



**Figure 4.** Mean scores of distal outcome variables assessed at one-year follow-up. *Note* Mean score with 95% confidence interval is presented. Only the above part of error bar is displayed to avoid visual clutter. Numbers over the zig-zag lines indicate the effect size (Cohen's  $d$ ) of pairwise comparisons.

Variable-frequency	Class 1	Class 2	Class 3	Class 4
Women-%	75.8%	79.6%	81.2%	73.8%
Age-M (years)	42.9	43.1	44.2	45.8
Highest education level <sup>a</sup>				
Primary	250 (10.7%)	534 (16.0%)	466 (13.7%)	224 (11.0%)
Secondary	1248 (53.6%)	1988 (59.4%)	1982 (58.4%)	1140 (55.9%)
Tertiary	829 (35.6%)	824 (24.6%)	945 (27.9%)	675 (33.1%)
Full-time employment-%	56.4%	39.3%	45.4%	58.8%
Visits to a pain clinic in the past year <sup>a</sup>				
0–1 times	310 (13.2%)	152 (4.5%)	214 (6.3%)	292 (14.0%)
2–3 times	682 (29.0%)	568 (16.7%)	767 (22.4%)	643 (30.8%)
≥ 4 times	1361 (57.8%)	2688 (78.9%)	2443 (71.3%)	1156 (55.3%)
Pain duration (years)-M (SD)	8.1 (9.1)	8.5 (9.2)	8.6 (9.0)	8.6 (9.9)
Social support-M (SD) <sup>b</sup>	3.74 (1.38)	4.18 (1.42)	4.40 (1.23)	4.14 (1.28)

**Table 1.** Descriptive statistics of background variables by classes. <sup>a</sup>Percentages show the proportions in each Class (i.e., column percentages). <sup>b</sup>The theoretical range of social support is between 0 and 6.

Variable	Class 1	Class 2	Class 3
Sex	-0.05	0.53***	0.67***
Age	-0.03***	-0.02***	-0.01***
Higher education level	0.15*	-0.27***	-0.18**
Full-time employment	-0.11	-0.79***	-0.60***
Visits to a pain clinic in the past year	0.10	0.93***	0.60***
Pain duration	0.00	0.01**	0.01**
Social support	-0.27***	0.00	0.20***

**Table 2.** The logit coefficients of predictors from the three-step approach. Class 4 is the reference group. With listwise deletion, 9173 cases (76.5%) with complete data were used for the above regression estimation. Sex: 1 = man, 2 = woman. Educational level: 1 = primary education, 2 = secondary education, 3 = tertiary education. Full-time employment: 1 = yes, 0 = no. Visits to a pain clinic in the past year: 0 = 0-1 time, 1 = 2-3 times, 2 = over 4 times. Pain duration: years of pain experience. Social support: a subscale from the Multidimensional Pain Inventory [29], with higher scores meaning more perceived social support for their pain problems. \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

less social support than Class 4, Class 3 members had more social support than Class 4 members. In addition, Class 3 members perceived more social support than Class 4 members, whereas Class 1 members perceived less.

## Discussion

Our study aimed to identify prognostically meaningful subgroups of chronic pain patients using data from a naturalistic and population level setting. In contrast to previous work that used unsupervised clustering approaches, a supervised machine learning framework was used for model selection and performance evaluation. Based on the optimal solution, four qualitatively distinct subgroups of chronic pain patients were discerned. The four-class solution showed a robust explanatory ability ( $R^2 = 26\%$ ) of the long-term follow-up on unseen data in the outer cross-validation, providing support for the predictive validity of derived subgroups using a rigorous approach for performance evaluation. This means that a significant part of the observed variation in pain outcome can be explained by a simple linear model of class-membership posterior probabilities estimated for each subject of the test-sets in our cross-validation. This was corroborated by effect size differences of non-trivial magnitude on pain interference and general health status at one-year follow-up between specific subgroups when the final model was fitted to the whole sample. The subgroups were also distinguishable based on examined background characteristics. Thus, as a novel method to combine an advanced unsupervised person-centered analytical method (i.e., LVMM) with a supervised learning approach to classification, our initial findings are encouraging, especially given that previous studies have generally failed to identify reliable predictors of long-term outcomes following interdisciplinary treatment using simpler regression- and variable-centered analytical approaches<sup>11</sup>.

In agreement with both mainstream recommendations and growing body of evidence<sup>27,28</sup>, our results show that psychological and pain characteristics (e.g., pain intensity) are both important phenotypic characteristics that interact at the level of the individual. The four subgroups identified in our study have their distinct characteristics. Class 4 (18.4%), with relatively low scores on pain characteristics, negative emotions, and better health status, appeared to be the healthiest subgroup. In contrast, with elevated levels of pain characteristics and negative emotions as well as poorer health status, Class 2 (29.8%) was an ill and sizable subgroup. Classes 1 (21.5%) and



3 (30.3%) showed a medium level on all indicators, although nuanced differences are observed between these two subgroups: Class 1 had higher anxiety, depression, and affective distress than Class 3, although members of this Class had less pain symptoms and better health status than Class 3.

These subgroups identified in our study have certain similarities with previous typologies of pain patients discovered in studies using various clustering methods. In a seminal study which also included both pain characteristics and psychological variables<sup>20</sup>, three subgroups among chronic pain patients were named as “dysfunctional”, “interpersonally distressed”, and “adaptive copers”. This three-subgroup taxonomy is comparable with Class 2, Class 1/3, and Class 4 in our study. Our findings are also in broad agreement with recent studies that have aimed to identify subgroups in patients receiving interdisciplinary treatment. Most relevant, a study using LVMM<sup>16</sup> also identified four similar subgroups of patients in the degree of pain severity, physical health, and emotional distress. While this study<sup>16</sup> relied on a conventional, exploratory approach to determining the number of subgroups, findings, taken together, support the four latent classes among chronic pain patients across samples.

A noteworthy finding, with potential clinical implications, is that two of the classes (Classes 1 and 2) had patients with marked emotional distress (anxiety and depression). As noted<sup>7</sup> after reviewing pain patient phenotype studies, research tends to discern a subgroup with greater pain intensity and negative emotions, highlighting the central phenotypic role of mental health status. Consistent with a substantial body of evidence supporting bi-directional links between general negative affect and severity of pain<sup>2</sup>, as well as with findings from previous studies on patient types<sup>6,57</sup>, increased negative emotions and pain presentations appear to cooccur. Specifically, the increased anxiety and depression observed in Class 2 (with the worse conditions across most indicators) highlight the key role of negative emotions among those individuals with worse clinical presentation at intake as well as increased risk of unfavorable long-term outcomes. In this context, the interrelationships between fear of movement, anxiety, and depression are also worth noting. A previous clustering study<sup>6</sup> found two clusters with similar levels of fear of movement, but one of these patient clusters showed distressed mood, as well as higher pain intensity.

The additive effect of emotional distress on both pain and fear may necessitate targeted interventions. Indeed, more recently, the importance of interventions targeting emotional processing and regulation have been highlighted for specific subgroups of chronic pain patients with high emotional distress<sup>58,59</sup>—treatment approaches that may be particularly valuable for Classes 1 and 2. In comparison to Class 1, Class 2 members showed elevated symptoms on all indicators, suggesting the need for a more comprehensive interdisciplinary treatment in targeted interventions. As previously recommended<sup>6</sup>, Class 2 members may also benefit specifically from exposure-based treatment given their high fear of movement. While specific treatment recommendations are in need of confirmatory tests, our findings call into question the one-size-fits-all approach to pain treatment and highlight the importance of identifying subgroups for improved treatment matching, which has long been recognized in the field<sup>5</sup>.

The associations between covariates and subgroup membership provide further information about the latent classes. A previous study<sup>16</sup> has identified a subgroup, which had lower scores on pain characteristics, but slightly more negative emotions; this subgroup members were remarkably older than others (10–15 years). Similar to this subgroup, Class 4 in our study, with less pain and negative emotions, was also older than other subgroups. It is important to note that Class 4 members also received more education and had less pain duration than those in Classes 2 and 3. An interesting contrast among groups was observed in social support. Compared to Class 4, the healthiest subgroup, Class 1 patients perceived less social support whereas Class 3 members perceived more. In previous subgroup studies, lack of social support has been found to be associated with being in subgroups which showed longer pain duration<sup>17</sup> or severer pain burden<sup>16</sup>. It may partially explain why Class 3 members in our study showed relatively lower levels of negative emotions, even though their pain intensity was similar to Class 2. In contrast, with lower social support, Class 1 members had more emotional distress than Classes 3 and 4, although its pain intensity was similar to Class 4. Our findings indicate that it is useful to consider the relationship between pain and emotional distress with the consideration of social support.

By combining the strengths of unsupervised and supervised learning approaches, our study also offers a novel methodological perspective on subgrouping in pain research. The framework expands on a recent methodological approach in which the selection of mixture model is governed by its prediction accuracy<sup>26</sup>. Using a nested cross-validation procedure, we were able to use an analgesic outcome in the selection of model as well as independently estimating the effects of the selected model for unseen cases. Replicability of clusters has always been a key focus because it indicates how reliable the cluster profiling is across samples<sup>20,60–62</sup>. However, failure to replicate different class solutions has been a source of concern, as has the ability to use subgroups for prognostic and treatment matching purposes. To date, simpler cross-validation procedures in which the sample data is split in to an exploratory and confirmatory part have been utilized to combat overfitting and replication issues in mixture modeling<sup>23</sup>. By combining a highly flexible unsupervised method with nested cross-validation and an optimization based on a supervised prediction of a one-year follow-up outcome, our study has continued this rigorous lineage.

While our statistical learning approach is promising for identifying prognostically meaningful subgroups, there are multiple ways to modify the analytical framework and widen its application. First, both the selection of indicators and the weights of chosen indicators could also have been optimized for outcome prediction using an appropriate supervised method prior to formation of latent classes in mixture modeling. Second, a single outcome variable was used as the target and a simple linear regression model was used to predict outcome from posterior class-membership probabilities; by using several outcomes and examining non-linear relations with outcomes could provide a more comprehensive evaluation of the models. Third, there are ways in which the finite mixture models could have been optimized further, including, for example, exploring additional individual heterogeneities within each class by combining mixture modeling with factor analysis (factor mixture model)—a method that has shown promise in identifying subgroups in specific pain conditions<sup>24</sup>. Finally, Gaussian mixture modeling

assumes normally distributed variables within each latent class and when this assumption is violated can create spurious class solutions. However, recent methodological developments have extended mixture model to handle non-normal data reducing the risk of extracting classes that are merely due to non-normality<sup>63</sup>. Nonetheless, a key strength of LVMM is that the framework enables empirical comparisons of a diverse set of models (e.g., growth mixture models, factor mixture models)—all done within one coherent analytical framework. Future research may build on this combined framework of unsupervised and supervised learning to incorporate information from additional sources (e.g., variables collected during treatment), examine individual change trajectories, and treatment response patterns in randomized experiments to identify clinically meaningful subgroups. Regardless of mixture model and applied context, the framework enables the identification of prognostically meaningful subgroups based on their ability to predict outcomes.

Our study has several limitations. The rehabilitation treatment varied across clinics and patients were not randomized to different treatments; therefore, the causes for outcome variations at follow-up remain unknown. Only patients with follow-up pain outcomes were included in this study; the reasons for non-adherence are not known, potentially biasing estimates due to non-ignorable causes of missing data. While guidelines about phenotyping pain patients are available, researchers have employed various strategies to select indicators. In our study, we focused on a restricted number of indicators and biological information such as genetic data and specific diagnosis were not included in the phenotypic characteristics. Moreover, social domains in our study are not fully represented. A considerable number of previous studies investigate subgroups among a specific type of pain patients (e.g., fibromyalgia, low back pain). Our study assumed that psychological aspects, self-perceived pain, and health status were independent of physical pathology across various types of pain conditions<sup>18</sup>. However, the assumption was not tested and nor did our study measure patients' beliefs about etiology, which have been related to recovery process<sup>64</sup>.

Notwithstanding, our study utilized a large, national sample follow-up assessments among chronic pain patients, mitigating limitations identified in previous pain patient classification studies. Using a rigors approach for identifying and evaluating our analytical model, our findings add to growing body of studies aiming to identify clinical meaningful subgroups across diverse chronic pain patients. From an applied methods standpoint, our study provides a novel person-centered supervised machine learning analytical framework that can be further expanded upon to ultimately provide an empirically derived taxonomy for treatment matching and prediction.

### Data availability

The datasets generated and/or analyzed during the current study are not publicly available due to privacy or ethical restrictions but are available from the corresponding author on reasonable request.

Received: 22 June 2023; Accepted: 17 May 2024

Published online: 31 May 2024

### References

- Hamburg, M. A. & Collins, F. S. The path to personalized medicine. *N. Engl. J. Med.* **363**, 301–304. <https://doi.org/10.1056/NEJMp1006304> (2010).
- Edwards, R. R. *et al.* Patient phenotyping in clinical trials of chronic pain treatments: IMMPACT recommendations. *PAIN Rep.* **6**, e896. <https://doi.org/10.1097/pr9.0000000000000896> (2021).
- Breivik, H., Collett, B., Ventafridda, V., Cohen, R. & Gallacher, D. Survey of chronic pain in Europe: Prevalence, impact on daily life, and treatment. *Eur. J. Pain* **10**, 287–287. <https://doi.org/10.1016/j.ejpain.2005.06.009> (2006).
- Gatchel, R. J., Peng, Y. B., Peters, M. L., Fuchs, P. N. & Turk, D. C. The biopsychosocial approach to chronic pain: Scientific advances and future directions. *Psychol. Bull.* **133**, 581–624. <https://doi.org/10.1037/0033-2909.133.4.581> (2007).
- Turk, D. C. The potential of treatment matching for subgroups of patients with chronic pain: Lumping versus splitting. *Clin. J. Pain* **21**, 44–55. <https://doi.org/10.1097/00002508-200501000-00006> (2005).
- Boersma, K. & Linton, S. J. Screening to identify patients at risk: Profiles of psychological risk factors for early intervention. *Clin. J. Pain* **21**, 38–43. <https://doi.org/10.1097/00002508-200501000-00005> (2005).
- Viniol, A. *et al.* Chronic low back pain patient groups in primary care—A cross sectional cluster analysis. *BMC Musculoskelet Disord.* **14**, 294. <https://doi.org/10.1186/1471-2474-14-294> (2013).
- Maixner, W., Fillingim, R. B., Williams, D. A., Smith, S. B. & Slade, G. D. Overlapping chronic pain conditions: Implications for diagnosis and classification. *J. Pain* **17**, T93–T107. <https://doi.org/10.1016/j.jpain.2016.06.002> (2016).
- Linton, S. J. & Shaw, W. S. Impact of psychological factors in the experience of pain. *Phys. Ther.* **91**, 700–711. <https://doi.org/10.2522/ptj.20100330> (2011).
- Gerdle, B. *et al.* Who benefits from multimodal rehabilitation—An exploration of pain, psychological distress, and life impacts in over 35,000 chronic pain patients identified in the Swedish Quality Registry for Pain Rehabilitation. *J. Pain Res.* **12**, 891–908. <https://doi.org/10.2147/JPR.S190003> (2019).
- Gerdle, B., Molander, P., Stenberg, G., Stålnacke, B.-M. & Enthoven, P. Weak outcome predictors of multimodal rehabilitation at one-year follow-up in patients with chronic pain—a practice based evidence study from two SQRP centres. *BMC Musculoskelet Disord.* **17**, 490. <https://doi.org/10.1186/s12891-016-1346-7> (2016).
- Khoury, M. J. & Ioannidis, J. P. A. Big data meets public health. *Science* **346**, 1054–1055. <https://doi.org/10.1126/science.aaa2709> (2014).
- Yim, Y.-R. *et al.* Identifying fibromyalgia subgroups using cluster analysis: Relationships with clinical variables. *Eur. J. Pain* **21**, 374–384. <https://doi.org/10.1002/ejp.935> (2017).
- Davis, F. *et al.* Characterizing classes of fibromyalgia within the continuum of central sensitization syndrome. *J. Pain Res.* **11**, 2551–2560. <https://doi.org/10.2147/JPR.S147199> (2018).
- Ringqvist, Å., Dragioti, E., Björk, M., Larsson, B. & Gerdle, B. Moderate and stable pain reductions as a result of interdisciplinary pain rehabilitation—A cohort study from the Swedish Quality Registry for Pain Rehabilitation (SQRP). *J. Clin. Med.* **8**, 905. <https://doi.org/10.3390/jcm8060905> (2019).
- Obbarius, A. *et al.* A step towards a better understanding of pain phenotypes: Latent class analysis in chronic pain patients receiving multimodal inpatient treatment. *J. Pain Res.* **13**, 1023–1038. <https://doi.org/10.2147/jpr.s223092> (2020).

17. Bäckryd, E., Persson, E. B., Larsson, A. I., Fischer, M. R. & Gerdle, B. Chronic pain patients can be classified into four groups: Clustering-based discriminant analysis of psychometric data from 4665 patients referred to a multidisciplinary pain centre (a SQRP study). *PLoS ONE* **13**, e0192623. <https://doi.org/10.1371/journal.pone.0192623> (2018).
18. Turk, D. C. & Okifuji, A. Psychological factors in chronic pain: Evolution and revolution. *J. Consult. Clin. Psychol.* **70**, 678–690. <https://doi.org/10.1037/0022-006X.70.3.678> (2002).
19. Pincus, T. & McCracken, L. M. Psychological factors and treatment opportunities in low back pain. *Best Pract. Res. Clin. Rheumatol.* **27**, 625–635. <https://doi.org/10.1016/j.berh.2013.09.010> (2013).
20. Turk, D. C. & Rudy, T. E. Toward an empirically derived taxonomy of chronic pain patients: Integration of psychological assessment data. *J. Consult. Clin. Psychol.* **56**, 233–238. <https://doi.org/10.1037/0022-006X.56.2.233> (1988).
21. Day, M. A., Ehde, D. M. & Jensen, M. P. Psychosocial pain management moderation: The limit, activate, and enhance model. *J. Pain* **16**, 947–960. <https://doi.org/10.1016/j.jpain.2015.07.003> (2015).
22. Clark, S. L. *et al.* Models and strategies for factor mixture analysis: An example concerning the structure underlying psychological disorders. *Struct. Equ. Model.* **20**, 681–703. <https://doi.org/10.1080/10705511.2013.824786> (2013).
23. Lubke, G. H. & Luningham, J. Fitting latent variable mixture models. *Behav. Res. Ther.* **98**, 91–102. <https://doi.org/10.1016/j.brat.2017.04.003> (2017).
24. Lipton, R. B. *et al.* Improving the classification of migraine subtypes: An empirical approach based on factor mixture models in the American Migraine Prevalence and Prevention (AMPP) study. *J. Headache Pain* **54**, 830–849. <https://doi.org/10.1111/head.12332> (2014).
25. Muthén, B. Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran (2003). *Psychol. Methods* **8**, 369–377. <https://doi.org/10.1037/1082-989X.8.3.369> (2003).
26. Jo, B. *et al.* Targeted use of growth mixture modeling: A learning perspective. *Stat. Med.* **36**, 671–686. <https://doi.org/10.1002/sim.7152> (2017).
27. Dworkin, R. H. *et al.* Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* **113**, 9–19. <https://doi.org/10.1016/j.pain.2004.09.012> (2005).
28. Kaiser, U. *et al.* Developing a core outcome domain set to assessing effectiveness of interdisciplinary multimodal pain therapy: The VAPAIN consensus statement on core outcome domains. *Pain* **159**, 673–683. <https://doi.org/10.1097/j.pain.0000000000001129> (2018).
29. Kerns, R. D., Turk, D. C. & Rudy, T. E. The west haven-Yale multidimensional pain inventory (WHYMPI). *Pain* **23**, 345–356. [https://doi.org/10.1016/0304-3959\(85\)90004-1](https://doi.org/10.1016/0304-3959(85)90004-1) (1985).
30. Nyberg, V. E., Novo, M. & Sjölund, B. H. Do multidimensional pain inventory scale score changes indicate risk of receiving sick leave benefits 1 year after a pain rehabilitation programme?. *Disabil. Rehabil.* **33**, 1548–1556. <https://doi.org/10.3109/09638288.2010.533815> (2011).
31. Gerdle, B., Fischer, M. R., Cervin, M. & Ringqvist, Å. Spreading of pain in patients with chronic pain is related to pain duration and clinical presentation and weakly associated with outcomes of interdisciplinary pain rehabilitation: A cohort study from the Swedish Quality Registry for Pain Rehabilitation (SQRP). *J. Pain Res.* **14**, 173–187. <https://doi.org/10.2147/JPR.S288638> (2021).
32. Zigmond, A. S. & Snaith, R. P. The hospital anxiety and depression scale. *Acta Psychiatr. Scand.* **67**, 361–370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x> (1983).
33. Lisspers, J., Nygren, A. & Söderman, E. Hospital anxiety and depression scale (HAD): Some psychometric data for a Swedish sample. *Acta Psychiatr. Scand.* **96**, 281–286. <https://doi.org/10.1111/j.1600-0447.1997.tb10164.x> (1997).
34. Vlaeyen, J. W. S., Crombez, G. & Linton, S. J. The fear-avoidance model of pain. *Pain* **157**, 1588–1589. <https://doi.org/10.1097/j.pain.0000000000000574> (2016).
35. Vlaeyen, J. W. S., Kole-Snijders, A. M. J., Boeren, R. G. B. & van Eek, H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain* **62**, 363–372. [https://doi.org/10.1016/0304-3959\(94\)00279-N](https://doi.org/10.1016/0304-3959(94)00279-N) (1995).
36. Lundberg, M. K. E., Styf, J. & Carlsson, S. G. A psychometric evaluation of the Tampa Scale for Kinesiophobia—from a physiotherapeutic perspective. *Physiother. Theory Pract.* **20**, 121–133. <https://doi.org/10.1080/09593980490453002> (2004).
37. Ware, J. E. J. & Sherbourne, C. D. The MOS 36-Item short-form health survey (SF-36): I. Conceptual framework and item selection. *Med. Care* **30**, 473–483 (1992).
38. Sullivan, M., Karlsson, J. & Ware, J. E. The Swedish SF-36 health survey—I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. *Soc. Sci. Med.* **41**, 1349–1358. [https://doi.org/10.1016/0272-9536\(95\)00125-Q](https://doi.org/10.1016/0272-9536(95)00125-Q) (1995).
39. Eccles, J. A. & Davies, K. A. The challenges of chronic pain and fatigue. *Clin. Med.* **21**, 19–27. <https://doi.org/10.7861/clinmed.2020-1009> (2021).
40. Ware, J. E., Snow, K. K., Kosinski, M. & Gandek, B. *SF-36 health survey: Manual and interpretation guide* (Quality Metric, 1993).
41. The EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* **16**, 199–208. [https://doi.org/10.1016/0168-8510\(90\)90421-9](https://doi.org/10.1016/0168-8510(90)90421-9)
42. Taylor, A. M. *et al.* Assessment of physical function and participation in chronic pain clinical trials: IMMPACT/OMERACT recommendations. *Pain* **157**, 1836–1850. <https://doi.org/10.1097/j.pain.0000000000000577> (2016).
43. Turk, D. C. *et al.* Identifying important outcome domains for chronic pain clinical trials: An IMMPACT survey of people with pain. *Pain* **137**, 276–285. <https://doi.org/10.1016/j.pain.2007.09.002> (2008).
44. McLachlan, G. & Peel, D. *Finite mixture models* (Wiley, 2000).
45. Muthén, B. & Shedden, K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* **55**, 463–469. <https://doi.org/10.1111/j.0006-341X.1999.00463.x> (1999).
46. Muthén, B. O. Beyond SEM: General latent variable modeling. *Behaviormetrika* **29**, 81–117. <https://doi.org/10.2333/bhmk.29.81> (2002).
47. Poldrack, R. A., Huckins, G. & Varoquaux, G. Establishment of best practices for evidence for prediction: A review. *JAMA Psychiatry* **77**, 534–540. <https://doi.org/10.1001/jamapsychiatry.2019.3671> (2020).
48. Danders, F. J. W. M., Traverso, A., Wee, L. & van Kuijk, S. M. J. Prediction modeling methodology. In *Fundamentals of clinical data science* (eds Kubben, P. *et al.*) 101–120 (Springer, 2019). [https://doi.org/10.1007/978-3-319-99713-1\\_8](https://doi.org/10.1007/978-3-319-99713-1_8).
49. Rodriguez, J. D., Perez, A. & Lozano, J. A. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**, 569–575. <https://doi.org/10.1109/TPAMI.2009.187> (2010).
50. Hosseini, M. *et al.* I tried a bunch of things: The dangers of unexpected overfitting in classification of brain data. *Neurosci. Biobehav. Rev.* **119**, 456–467. <https://doi.org/10.1016/j.neubiorev.2020.09.036> (2020).
51. Stone, M. Cross-validated choice and assessment of statistical predictions. *J. R. Stat. Soc. Ser. B Methodol.* **36**, 111–133. <https://doi.org/10.1111/j.2517-6161.1974.tb00994.x> (1974).
52. Cawley, G. C. & Talbot, N. L. C. On over-fitting in model selection and subsequent selection bias in performance evaluation. *JMLR* **11**, 2079–2107 (2010).
53. Hallquist, M. N. & Wiley, J. F. MplusAutomation: An R package for facilitating large-scale latent variable analyses in Mplus. *Struct. Equ. Model. Multidiscip. J.* **25**, 621–638. <https://doi.org/10.1080/10705511.2017.1402334> (2018).
54. Muthén, L. K. & Muthén, B. O. *Mplus user's guide*. 8 edn, (Muthén & Muthén, 1998–2017).
55. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **28**, 1–26. <https://doi.org/10.18637/jss.v028.i05> (2008).

56. Asparouhov, T. & Muthén, B. Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Struct. Equ. Model.* **21**, 329–341. <https://doi.org/10.1080/10705511.2014.915181> (2014).
57. Santana, A. N., de Santana, C. N. & Montoya, P. Chronic pain diagnosis using machine learning, questionnaires, and QST: A sensitivity experiment. *Diagnostics* **10**, 958. <https://doi.org/10.3390/diagnostics10110958> (2020).
58. Lumley, M. A. *et al.* Emotional awareness and other emotional processes: Implications for the assessment and treatment of chronic pain. *Pain Manag.* **11**, 325–332. <https://doi.org/10.2217/pmt-2020-0081> (2021).
59. Boersma, K. *et al.* Efficacy of a transdiagnostic emotion-focused exposure treatment for chronic pain patients with comorbid anxiety and depression: A randomized controlled trial. *Pain* **160**, 1708–1718. <https://doi.org/10.1097/j.pain.0000000000001575> (2019).
60. Jamison, R. N., Rudy, T. E., Penzien, D. B. & Mosley, T. H. J. Cognitive-behavioral classifications of chronic pain: Replication and extension of empirically derived patient profiles. *Pain* **57**, 277–292. [https://doi.org/10.1016/0304-3959\(94\)90003-5](https://doi.org/10.1016/0304-3959(94)90003-5) (1994).
61. Wilson, H. D., Robinson, J. P. & Turk, D. C. Toward the identification of symptom patterns in people with fibromyalgia. *Arthritis Care Res.* **61**, 527–534. <https://doi.org/10.1002/art.24163> (2009).
62. Wilson, M. W., Richards, J. S., Klapow, J. C., DeVivo, M. J. & Greene, P. Cluster analysis and chronic pain: An empirical classification of pain subgroups in a spinal cord injury sample. *Rehabil. Psychol.* **50**, 381–388. <https://doi.org/10.1037/0090-5550.50.4.381> (2005).
63. Muthén, B. & Asparouhov, T. Growth mixture modeling with non-normal distributions. *Stat. Med.* **34**, 1041–1058. <https://doi.org/10.1002/sim.6388> (2015).
64. DeGood, D. E. & Kiernan, B. Perception of fault in patients with chronic pain. *Pain* **64**, 153–159. [https://doi.org/10.1016/0304-3959\(95\)00090-9](https://doi.org/10.1016/0304-3959(95)00090-9) (1996).

### Author contributions

H.H. conceived the study. H.H., X.Z., and D.R. designed the study. K.D., X.Z., D.R., and H.H. analyzed data. X.Z. and H.H. co-wrote the first draft, which all authors revised for critical content. All authors approved the final manuscript.

### Funding

Open access funding provided by Örebro University. This study was funded by a grant from the AFA insurance (AFA försäkringar DNR 190054). The authors would like to thank the DU/Pain group at Dalarna University for sharing data from their research database FRIDA. FRIDA is funded by Swedish Research Council (Vetenskapsrådet: 2015-02512) and by the Swedish Research Council for Health, Working Life and Welfare (FORTE: 2017-00177).

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-62542-w>.

**Correspondence** and requests for materials should be addressed to H.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024