An Ensemble of Psychological and Physical Health Indices Discriminates Between Individuals with Chronic Pain and Healthy Controls with High Reliability: A Machine Learning Study

Linda A. Antonucci ⬤ · Alessandro Taurino · Domenico Laera ·

Paolo Taurisano · Jolanda Losole · Sara Lutricuso · Chiara Abbatantuono ·

Mariateresa Giglio · Maria Fara De Caro · Giustino Varrassi ·

Filomena Puntillo

ABSTRACT

Introduction: Chronic pain (CP) is a complex multidimensional experience severely affecting individuals' quality of life. Multiple cognitive, affective, emotional, and interpersonal factors play a major role in CP. Furthermore, the psychological, social, and physical circumstances leading to CP show high inter-individual variability, thus making it difficult to identify core syndrome characteristics. In a biopsychosocial perspective, we aim at identifying a pattern of psycho-physical impairments that can reliably discriminate between CP individuals and healthy controls (HC) with high accuracy and estimated generalizability using machine learning.

Methods: A total of 118 CP and 86 HC were recruited. All individuals were administered several scales assessing quality of life, physical and mental health, personal functioning, anxiety, depression, beliefs about medical treatments, and cognitive ability. These features were trained to separate CP from HC using support vector classification and repeated nested cross-validation.

Results: Our psycho-physical classifier was able to discriminate CP from HC with 86.5% balanced accuracy and significance ($p = 0.0001$). The most reliable features characterizing CP were anxiety and depression scores, and belief of harm from prolonged pharmacological treatments; for HP, the most reliable features were physical and occupational functioning, and vitality levels.

Conclusion: Our findings suggest that, using psychological and physical assessments, it is possible to classify CP from HC with high reliability and estimated generalizability via (i) a pattern of psychological symptoms and cognitive beliefs characteristic of CP, and (ii) a pattern of intact physical functioning characteristic of HC. We think that our algorithm enables novel insights into potential individualized targets for CP-related early intervention programs.

L. A. Antonucci (✉) · A. Taurino · D. Laera
Department of Education, Psychology and Communication Science, University of Bari Aldo Moro, Bari, Italy
e-mail: linda.antonucci@uniba.it

P. Taurisano · J. Losole · S. Lutricuso ·
C. Abbatantuono · M. F. De Caro
Department of Basic Medical Science, Neuroscience and Sense Organs, University of Bari Aldo Moro, Bari, Italy

M. Giglio · F. Puntillo
Pain Therapy Unit, Department of Interdisciplinary Medicine, University of Bari Aldo Moro, Bari, Italy

G. Varrassi
Paolo Procacci Foundation, Rome, Italy

Key Summary Points

Why carry out this study?

Chronic pain (CP) is a complex multidimensional experience severely affecting the quality of life of individuals. Multiple cognitive, affective, emotional, and interpersonal factors play a major role in CP.

The psychological, social, and physical circumstances leading to CP show high inter-individual variability, thus making it difficult to identify core syndrome characteristics.

In a biopsychosocial perspective, we aim at identifying a pattern of psycho-physical impairments that can reliably discriminate between CP individuals and healthy controls (HC) with high accuracy and estimated generalizability using machine learning.

What was learned from the study?

Our psycho-physical classifier could discriminate CP from HC with 86.5% balanced accuracy and significance ($p = 0.0001$). The most reliable features characterizing CP were anxiety and depression scores, and belief of harm consequent to prolonged pharmacological treatments; for HP, the most reliable features were physical and occupational functioning, and vitality levels.

We think that our algorithm provides novel insights about potential individualized targets for CP-related early intervention programs. We think that our algorithm provides novel insights about potential individualized targets for CP-related early intervention programs.

INTRODUCTION

Pain is defined as an unpleased sensory or emotional experience associated with actual or potential tissue damage, or described in terms of such damage [1, 2]. While acute pain is an early warning sign that usually elicits reflex withdrawal and thereby promotes survival [3], chronic pain (CP) lasts longer than the injury itself, usually more than 3 months, and is an expression of nervous system maladaptation. CP affects 10–30% of the adult population in Europe [4, 5]. It is nowadays a major issue for public health and is associated with high economic and social burden [6], not only for individuals experiencing pain but also for their family, caregivers, and friends [7].

CP was recently defined as "a complex multidimensional experience" [4], to the extent that the biopsychosocial model is now considered the most valid paradigm for understanding the complexity of this syndrome. Indeed, in addition to physical functioning impairments, CP individuals show decreased productivity, increased mood swings or more severe depressive mood, impaired social functioning, increased sleep disturbances, decreased cognitive ability [8], and decreased participation in leisure activities [9] compared with healthy controls (HC). Overall, these impairments lead to a marked decrease in the overall perceived and objective quality of life in people experiencing CP, compared with HC. Notably, not only is CP architecture complex and multifactorial, but CP individuals are also very heterogeneous. Furthermore, the psychological, social, and physical circumstances leading to CP show a high degree of inter-individual variability [10]. Therefore, the identification of a psycho-physical signature able to reliably discriminate CP from HC at the single-subject level [11] may aid in diagnostic categorization. Such a signature could potentially provide novel insights about specific targets that might be central in individualized early identification and intervention programs.

So far, studies have characterized CP and HC differences in terms of group differences through univariate statistics, which are limited in terms of generalizability assessments [12, 13]. Moreover, to understand whether a syndrome-associated characteristic could also be qualified as a marker (i.e., as a measurable feature associated with a certain condition or process [14, 15]), one should investigate its sensitivity and specificity in identifying the respective patient population [12]. A promising way of addressing this question is by employing machine learning, which allows one to quantify the sensitivity, specificity, and generalizability of a disease signature at the single-subject level [16, 17], rather than just describing it at the group level. Therefore, machine learning allows one to (i) grasp the complex architecture of CP given by the multiple variables associated with the condition, and (ii) potentially clarify the factors underlying the high inter-individual heterogeneity between CP individuals.

While machine learning techniques have been recently applied to CP classification using neuroimaging data [18–21] to the best of our knowledge no study has employed self-administered or observer-rated clinical assessments within a machine learning environment with the same aim. Therefore, following a biopsychosocial perspective, the aim of this study is to identify a pattern of psychological and physical impairments that can reliably discriminate between CP and HC with high accuracy and estimated generalizability using machine learning.

METHODS

Sample Determination

A total of 118 consecutive outpatients (63.6% female; mean age 57.1 years) with chronic pain (CP), treated with analgesic medication, were recruited from the Pain Therapy Clinic, University of Bari General Hospital. Individuals diagnosed with CP had a subjective experience of physical pain, localized anywhere in the body, for at least 3 months, as defined by the European Federation of International

Association for the Study of Pain (IASP). In addition, 86 healthy controls (HC) were recruited either through flyers or among the CP individuals' and caregivers' (i.e., those who accompanied CP individuals to the visit) social networks. Therefore, our HC group was mainly composed of individuals who (i) were in contact with either CP individuals or, if present and available, their caregivers, (ii) learned of the research study via either CP individuals or, if present and available, their caregivers, and (iii) were willing to take part in the research. No CP caregiver entered the study as a HC. Common exclusion criteria for HC and CP were the presence of any other significant clinical condition (besides the CP diagnosis) or the presence of any Axis I psychiatric disorder, according to the Structured Clinical Interview for DSM-IV [22]. The demographic and clinical characteristics of CP and HC individuals are reported in Table 1. Independent-samples t tests, chi-square tests, and Fisher's exact tests were employed to investigate the presence of any demographic differences between the two groups. All p values were multiple-comparison-corrected through the false discovery rate (FDR) method, following published procedures [23]. The significance was determined at $\alpha = 0.05$.

This research was conducted ethically in accordance with the World Medical Association (www.wma.net) Declaration of Helsinki of 1964 and its later amendments. All individuals have given their written informed consent before entering the study. The research protocol has been approved by Bari University Hospital local ethical committee.

Psychological Assessment

All individuals were administered the following questionnaires:

1. Beliefs about Medicines Questionnaire—General section (BMQ-General), to broadly investigate cognitive beliefs about the utility and the harm of pharmacological treatments; the BMQ-General section includes two factors: the evaluation of to what degree medicines are harmful and agreement with the fact that they should not be

Table 1 Demographic characteristics of chronic pain (CP) and healthy control (HC) individuals

	CP	HC	t/Chi-square/Fisher's exact test, effect size, p value	Multiple-comparison-corrected p value
Age in years, mean (SD)	57.07 (11.54)	57.12 (13.29)	−0.027, 0.004, 0.978	0.978
Gender ratio, M/F	75/43	56/30	0.05, 0.032, 0.818	0.978
Education, mean (SD)	9.47 (3.81)	9.92 (3.96)	−0.803. 0.115, 0.423	0.705
Marital status, no. (%)	Divorced = 10 (8%)	Divorced = 5 (6%)	**9.792, 0.01, 0.044**	0.11
	Married = 80 (68%)	Married = 68 (79%)		
	Widower = 13 (11%)	Widower = 6 (7%)		
	Single = 13 (11%)	Single = 2 (3%)		
	Living with partner = 2 (2%)	Living with partner = 5 (6%)		
Occupational status, no. (%)	Retired = 40 (34%)	Retired = 24 (28%)	**50.533, 0.01, 0.001**	**0.001**
	Household = 33 (30%)	Household = 21 (24%)		
	Employee = 4 (3%)	Employee = 9 (10%)		
	Freelance = 0 (0%)	Freelance = 19 (23%)		
	Unemployed = 20 (16%)	Unemployed = 1 (1%)		
	Workmen = 14 (11%)	Workmen = 2 (2%)		
	Fired = 7 (6%)	Fired = 10 (12%)		

Significant differences between CP and HC ($p < 0.05$) are marked in bold

taken continuously (General-Harm), and agreement with the fact that doctors tend to prescribe too many drugs (General-Overuse). The BMQ-Specific section was administered to individuals but was purposely excluded from further analyses in order to avoid introducing in the experiment a bias related to the fact that CP individuals were under analgesic treatment, while HC were not. The Italian version of the BMQ shows high validity and reliability properties [24].

2. The Hospital Anxiety and Depression Scale (HADS), to determine individual levels of anxiety and depression; HADS Italian

version shows good validity and reliability [25].

3. 36-Item Short Form Health Survey (SF-36), to evaluate the individual levels of quality of life across seven domains: vitality, physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning, social role functioning, mental health. For analysis purposes, the "bodily pain" domain was excluded from further analyses given that this factor could be highly related to the label (i.e., CP vs. HC). The Italian version of the SF-36 [26] has high reliability and good validity.

4. Coloured Progressive Matrices (CPM), to evaluate non-verbal intelligence. Both raw and age- and education-corrected scores were entered in the analyses. CPM have very good psychometric properties, especially in terms of validity and reliability [27].

Machine Learning Analysis

The overall analytical strategy was to use all the subscales and scores acquired through the BMQ-General, HADS, SF-36 (bodily pain subscale excluded, see Sect. 2.2), and the CPM to build a psycho-physical multivariate model able to accurately discriminate between HC and CP. In total, the model was built based on 13 features, which are all reported in Table 2. Two-sample t tests were employed to assess differences between HC and CP for each of the features entering the machine learning algorithm. All p values were < 0.05, FDR corrected for all the subscales used [23]. Machine learning analyses were performed using NeuroMiner version 1.0 software (https://github.com/neurominer-git?tab=repositories). All analysis steps are described in the following sections.

Cross-Validation Framework
To prevent information leaking between individuals used for training and testing the models [28], we built a double cycle, nested cross-validation (CV) framework [16]. Indeed, we split the data first into training and test sets on an

outer (CV2) cycle, and then we split the resulting training folds again into an inner (CV1) training and test data cycle [29]. Therefore, nested CV induces a strict separation between training and test data. This way, in a machine learning framework, parameter optimization is performed within the inner (CV1) cycle, and generalization error estimation is performed only from the outer (CV2) cycle. CV2 samples never visited the classification algorithms during the entire training process [28]. In both inner (CV1) and outer (CV2) CV levels, we employed a tenfold CV cycle. We extended nested CV to repeated nested CV [12] at both the inner and outer cross-validation cycles by randomly permuting the participants within their groups (number of permutations: in CV1 = 5, in CV2 = 10) and repeating the CV cycle for each of these permutations.

Data Preprocessing
Our NeuroMiner machine learning preprocessing pipeline consisted of the following steps:

1. As many machine learning algorithms are sensitive to scale differences between features, we scaled each variable to a 0–1 range to remove these effects from each training sample matrix. The scaling parameters were then applied to the inner and outer CV.

2. To avoid the effect of any demographic confounds on the algorithm performance, features' scaled scores were further preprocessed through correction for age, gender, education, work, and marital condition. Specifically, we removed the variance associated with these demographic variables within each inner and outer CV fold through partial correlations.

Feature Selection and Machine Learning Algorithm Implementation
Features included in the algorithm underwent a stepwise forward variable selection process [30] using a linear support vector machine (SVM) [31]. Specifically, data entered a greedy forward search wrapper [30] which allows for the identification of the most parsimonious subset of variables within the given variable pool, thus

Table 2 Chronic pain (CP) and healthy controls (HC) mean and standard deviation values for each of the features entered in the machine learning algorithm (assessments are fully described in Sect. 2.2)

	CP	HC	t, effect size, p value	Multiple-comparison-corrected p value
BMQ harm, mean (SD)	12.92 (2.88)	11.52 (2.99)	**3.348, 0.476, 0.001**	**0.001**
BMQ overuse, mean (SD)	12.53 (3.19)	12.97 (2.60)	−1.081, 0.151, 0.281	0.281
HADS anxiety, mean (SD)	11.31 (6.16)	5.65 (4.59)	**7.506, 1.041, 0.001**	**0.001**
HADS depression, mean (SD)	9.69 (4.93)	6.01 (3.69)	**6.086, 0.845, 0.001**	**0.001**
SF-36 physical functioning, mean (SD)	48.56 (25.62)	86.05 (16.21)	**−12.767, 1.748, 0.001**	**0. 001**
SF-36 role physical functioning, mean (SD)	10.81 (23.90)	70.93 (33.83)	**−14.111, 2.052, 0.001**	**0.001**
SF-36 general health perceptions, mean (SD)	35.37 (18.46)	57.70 (14.38)	**−9.702, 1.349, 0.001**	**0.001**
SF-36 vitality, mean (SD)	37.25 (22.32)	58.90 (13.35)	**−8.625, 1.177, 0.001**	**0.001**
SF-36 social role functioning, mean (SD)	49.97 (29.02)	72.72 (15.77)	**−7.181, 0.974, 0.001**	**0.001**
SF-36 emotional role functioning, mean (SD)	26.51 (35.77)	74.67 (35.85)	**−9.484, 1.344, 0.001**	**0.001**
SF-36 mental health, mean (SD)	43.02 (23.18)	64.33 (13.81)	**−8.185, 1.117, 0.001**	**0.001**
CPM raw score, mean (SD)	24.88 (7.05)	27.23 (5.14)	**−2.746, 0.429, 0.007**	**0.007**
CPM corrected score, mean (SD)	2.45 (1.56)	3.02 (1.25)	**−2.918, 0.404, 0.004**	**0.004**

Significant differences between CP and HC ($p < 0.05$) are marked in bold

providing maximum prognostic performance with the smallest amount of predictive features. The wrapper algorithm used an SVM to evaluate the predictive value of each feature, then extracted the most predictive variable and reiterated over the remaining variable pool to select the second best performing variable, which was added to the first one. This process was reiterated until the optimal variable subspace had been identified. We stopped the variable search when the top 20% of the variables had been extracted by the wrapper, thus allowing us to identify a clinically applicable set of top-performing variables for classification purposes.

The wrapper-based feature selection was carried out for each CV1 training and test sample and then repeated for every combination of the SVM parameters C (misclassification

cost) and γ (kernel width) within a grid defined by the ranges $C = [0.0156—16]$ and $\gamma = [3.0518^{-5}—8]$. In each variable evaluation step in the CV1, the SVM algorithm modeled linear relationships between features and classification labels (HC vs. CP). In the linear kernel space, the SVM optimized a hyperplane that maximized separability between most HC-like and most CP-like subjects (i.e., the support vectors). Based on the trained hyperplane, the algorithm then predicted subjects' classification (HC vs. CP) of the inner CV1 cycle by projecting its data into the CV2 learned kernel space and measuring their geometric distance to the decision boundary. This resulted in a decision value and a predicted classification label per participant.

Investigation of Individual Features' Relevance Within the Machine Learning Algorithm

To better understand which variables might inform CP and HC classes at the single-subject level, we checked which features were the most reliable. Reliability for each feature is defined in terms of a cross-validation ratio (CVR = mean(w)/standard error(w)) [32]. In this formula, w represents the normalized individual weights from SVM models generated in the repeated nested CV scheme. Normalization is performed using the Euclidean norm of w, defined as $s = w/||w||2$ [32]. A positive CVR for each feature indicates higher CVR scores in CP compared to HC, while a negative CVR for each feature indicates higher CVR scores for HC compared to CP.

Permutation Testing

To assign statistical significance to the observed classification performance, we employed permutation [31]. We performed 1000 random permutations of the outcome labels (ie, HC vs. CP). For each permutation, we retrained all linear SVM models in the repeated nested CV experiment using the respective feature subsets obtained from the observed-label analyses. For each permutation, we accumulated the predictions of the random models into a permuted ensemble prediction for each outer cycle subject. Thus, we built a null distribution of out-of-

training classification performance (BAC) for every unimodal classifier. Finally, we calculated the significance of the observed out-of-training BAC as the number of events where the permuted out-of-training BAC was higher than or equal to the observed BAC divided by the number of permutations performed. The significance of the model was determined at $\alpha = 0.05$.

RESULTS

Demographic Differences Between Samples

HC and CP individuals did not differ by age, gender, or education level (all $p > 0.05$, Table 1). However, marital status and occupational status were differentially distributed across HC and CP individuals, respectively (both $p = 0.001$, Table 1). With regard to the features that entered the machine learning algorithm, CP had higher HADS scores on both anxiety and depression subscales, in BMQ harm (all $p = 0.001$, Table 2), and CPM raw and corrected scores (respectively, $p = 0.007$ and $p = 0.004$, Table 2); on the other hand, HC had higher scores than CP in all the SF-36 subscales entering the algorithm (all $p = 0.001$, Table 2).

Machine Learning Results

The cognitive classifier correctly discriminated CP from HC with a cross-validated balanced accuracy (BAC) of 86.5% and was significant at $p < 0.001$, with an area under the curve (AUC) of 0.92. Detailed classification metrics are reported in Table 3. Out of all 13 features originally included in the model, those with the highest positive CVR were HADS-depression, HADS-anxiety, and BMQ-overuse, while those with the highest negative CVR were SF-36 vitality, physical functioning, and physical role functioning (Table 4, Fig. 1).

Table 3 Validated classification performance of the classifier trained based on psycho-physical assessments within a repeated nested cross-validation framework

	True negatives	True positives	False negatives	False positives	Sensitivity	Specificity	Balanced accuracy	Positive predictive value	Negative predictive value	Number needed to diagnose	Positive likelihood ratio	Diagnostic odds ratio	Permutation test, p value	Youden's J statistic
HC vs. CP classification	105	72	14	13	88.2	84.7	86.5	89.0	83.7	1.4	5.8	0.1	0.001	0.7

Table 4 Cross-validation ratio (CVR) score of each feature within the machine learning algorithm, representing its reliability

Feature name	CVR score
HADS depression	0.99
BMQ harm	0.99
HADS anxiety	0.97
BMQ overuse	0.8
CPM raw score	0.51
CPM corrected score	0.34
SF-36 social role functioning	−0.4
SF-36 emotional role functioning	−0.54
SF-36 mental health	−0.81
SF-36 general health perceptions	−0.83
SF-36 vitality	−0.94
SF-36 physical functioning	−1
SF-36 role physical functioning	−1

A positive CVR for each feature indicates higher scores in chronic pain (CP) individuals compared to healthy controls (HC), while a negative CVR for each feature indicates higher scores for HC compared to CP

DISCUSSION

In the current study, we aimed at building a multivariate classification model through machine learning techniques that was able to discriminate between HC and CP individuals. In a biopsychosocial perspective, we built this classification model based on a wide variety of psychological, physical, and mental health-related features, as well as cognitive features, while strictly controlling for any potential demographic confounds (age, gender, education, marital status, and occupational status).

Before entering the machine learning framework, we observed that CP had higher HADS and BMQ harm subscales scores compared with HC. Concerning HADS, the higher depression scores we found in CP individuals compared with HC are consistent with a large body of literature demonstrating that CP
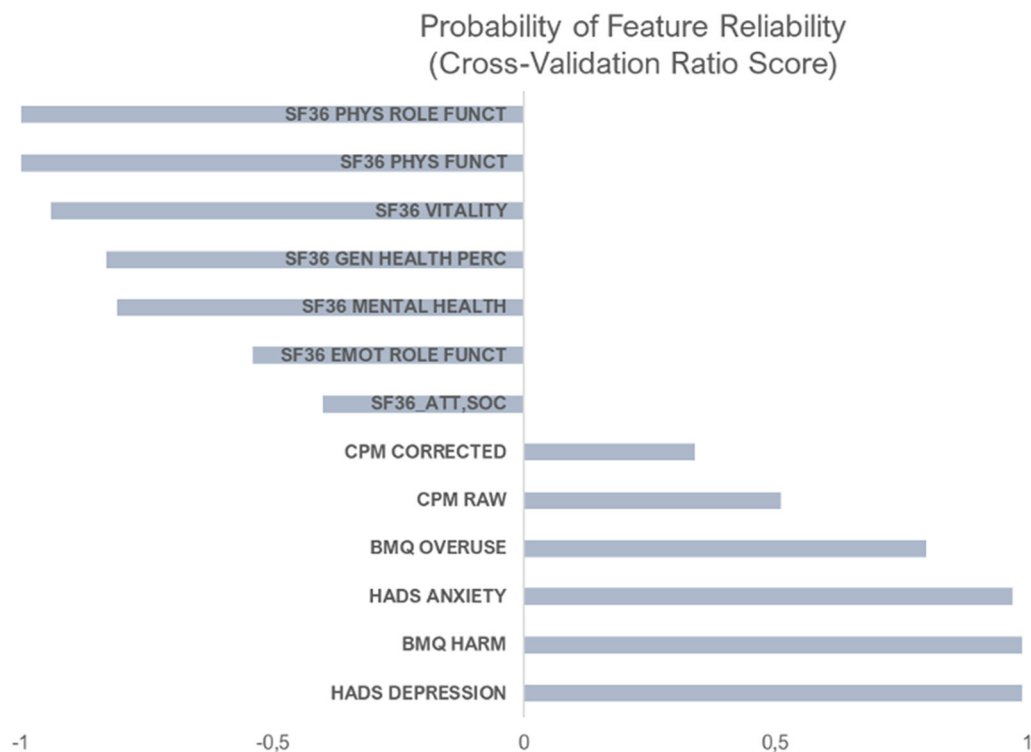
Fig. 1 Depiction of the cross-validation ratio scores, representing the reliability of each feature included the algorithm. A positive CVR for each feature indicates higher CVR scores in chronic pain (CP) individuals compared to healthy controls (HC), while a negative CVR for each feature indicates higher CVR scores for HC compared to CP

individuals very often develop depression after the CP diagnosis [33], to the extent that approximately 85% of patients with CP suffer from severe depression [34]. On the other hand, the higher anxiety scores in CP compared with HC match previous findings elucidating that fear and anxiety levels in CP individuals are associated with greater pain-related perceived disability via avoidance, cognitive preoccupation, and stress-related muscle activity [10]. Higher BMQ scores in CP compared with HC may reflect the fact that CP individuals more strongly believe that medicines may be harmful [34]. Although CP individuals were under analgesic medication, this finding should not be affected by ongoing treatment, given that the BMQ-General section is designed to capture overall perceptions of medication in general and despite any potential ongoing chronic illness (in contrast to BMQ-Specific, which indeed

was not included in the machine learning algorithm).

Our results also reveal that HC have higher CPM scores than CP, thus showing higher overall cognitive ability. This finding is consistent with several previous studies (for a review, see [2]). However, recent views have proposed that CP may be not directly associated with cognition itself but may potentially influence it via the other comorbidities that are frequently associated with CP, such as anxiety, depression, or emotional dysregulation. If validated, this hypothesis would imply future challenges for new targets within CP early intervention programs [8].

Furthermore, in all the SF-36 sub-dimensions, we found that all HC had higher scores than CP. Indeed, the CP state is very often associated with reduced physical activity [4], and this reduction is associated with the intensity, duration, and location of pain. This,

in turn, seems to affect patients' overall quality of life [35], such as emotional, psychological, and social functioning, which indeed were found to be lower in our CP individuals compared with HC. Of note, the smaller CP-HC difference across the SF-36 sub-dimensions was found for the "general health perceptions" subscale. Indeed, previous studies found that CP individuals are often not -conscious of the degree of physical impairment experienced and tend to overestimate their ability and do not feel impaired [4, 36]. This may have potential implications for CP early intervention strategies, since making patients more conscious of their impairments and behaviors can potentially promote a healthier and more active lifestyle.

Our machine learning results show for the first time that a pattern of psychological and physical health indices can discriminate between CP and HC with high balanced accuracy and significance. Indeed, when pooling a set of variables clearly associated with the CP experience based on univariate studies [4], and using them to discriminate CP from HC, we are able to correctly discriminate a CP individual from a HC at the single-subject level in 86.5% of cases within our repeated-nested cross-validation framework. Notably, as in previous studies using the same SVM technique [11, 15], we employed a stringent separation of training and test sets, and a robust, repeated-nested CV scheme. These methodological choices are in line with recent recommendations (for a review, see [12]), which noted that the gold-standard CV scheme ensuring the highest degree of reliability and generalizability of machine learning findings, in the absence of external replication samples, is nested CV. Thus, the psycho-physical HC versus CP classification model we have developed not only is highly accurate but also shows a very good extent of estimated generalizability.

The high accuracy and estimated generalizability our SVM algorithm further validate our biopsychosocial model's application in clinical practice, as it demonstrates (1) how several sensory, cognitive/affective, and interpersonal factors together contribute to CP syndrome, and (2) that CP is associated with several

psychological and physical processes that, in turn, affect the pain experience [36]). Specifically, we observed that, within this pool of variables clearly associated with the CP experience based on univariate studies [4], not all the features have the same discriminatory power; some are highly characteristic of only the CP status, and others of only the HC status. Indeed, within our algorithm, higher HADS-depression, HADS-anxiety, and BMQ-overuse scores are the most relevant features in discriminating a CP from a HC, thus showing higher sensitivity potential, while higher SF-36 vitality, physical functioning, and physical role functioning scores are the most reliable features for discriminating a HC from a CP, thus showing higher specificity potential. HADS findings further confirm the tight link between CP and depression. Indeed, depression and CP seem to influence each other, their respective development, and their respective severity [33]. The prognostic relevance of depression for CP is further supported by recent views highlighting that depressed CP individuals have a poorer prognosis than non-depressed CP [36]. On the other hand, CP-related anxiety has proven to influence the pain experience in multiple ways. Indeed, pain may cause feelings of anxiety, which in turn may increase individual pain sensitivity and make the experience of pain more persistent [37]. Furthermore, anxiety and CP share common cognitive and behavioral processes, such as increased attention towards threatening stimuli and avoidance of physical exertion [38]. On the prognostic level, pain-related fear and anxiety have been previously associated with greater disability and persistent pain experience [39]. The relevant role of the BMQ-overuse subscale in our HC-CP psycho-physical algorithm as a psychological feature which is more "CP-like" than "HC-like" is also not surprising. Indeed, as recently proposed [34], CP individuals may consider that their condition is irreversible and that, regardless of whether any therapy could be helpful, a relapse might always occur. This view is also coherent with the increased attention towards threats characteristic of CP individuals [38].

On the other hand, findings revealed that the features contributing the most to the

model's accuracy being more prototypical of HC compared with CP were those related to physical functioning, occupational functioning, and vitality (i.e., perceived energy and fatigue). These findings further confirm that, compared with HC, who don't show any significant limitations in these areas, the physical and occupational functioning of CP and their perceived energy are significantly affected by the pain experience itself. Previous studies have shown that pain is significantly associated with both psychological and physical dimensions of health-related quality of life [4], and that in each health-related quality of life dimension, CP score significantly worse than HC [40]. However, consistent with our findings, the greater impact of pain is on physical, rather than on mental, quality of life indices [41]. As for occupational functioning, their relevance in our SVM algorithm is coherent with the fact that, differently from HC, CP individuals often experience difficulties in their workplaces. Indeed, CP is often associated with higher absenteeism, early retirement, and more days of sick leave [42], especially for back pain and rheumatic diseases [4]. Furthermore, CP individuals are often forced to change their duties at their workplace due to their physical and psychological symptoms, and this may result in the loss of their jobs [41].

Taking these findings together, it seems that our algorithm identified a pattern of psychological feelings, disbeliefs, and cognitive distortions that are highly characteristic of CP individuals and reliably discriminate them from HC. On the other hand, it identifies a pattern of physical, occupational, and energy-related features highly characteristic of HC, and significantly less present in CP individuals.

Limitations

This study has certain limitations. Despite the stringent training and test data separation in our SVM algorithm and the nested CV employed, replication of our findings in independent and geographically different samples is needed to ensure that our findings are replicable, as external validation is considered the gold

standard in the field for assessing a model's effective (and not just estimated) generalizability. Moreover, it should be noted that our HC sample was composed mainly of individuals linked to the social circle of CP individuals' caregivers. Although this allowed us to build a HC sample with similar age and education level relative to those of our CP sample, it does not guarantee a purely random sampling of the HC group. Larger and randomly sampled external validation groups are needed to further test the generalizability of our psycho-physical machine learning algorithm. Another limitation that should be taken into account is that further clinically relevant aspects of the pain experience such as pain severity, duration of drug treatment, or history of treatments tried were not considered in this study. If on the one hand this is probably a consequence of the employment of broad inclusion/exclusion criteria, on the other hand, the general aim of the research project was to include in the study CP individuals based solely on diagnosis, and irrespective of their past pain experiences. Nevertheless, we think that future studies investigating the potential association between the single-subject-level decision scores generated through our machine learning algorithm and these highly relevant clinical aspects of the pain experience are warranted to provide deeper insights into the potential translation into clinical practice of our psycho-physical classification model. Consistently, our neuropsychological findings are limited by the fact that in our study, the cognitive area is represented solely by the CPM test. Although beyond the scope of this study, a better characterization of individual neurocognitive functions through a broader battery assessing specific (rather than general) cognitive functioning and sub-domains would have been more informative about the potential existence of specific neuropsychological assessments that could capture core pain-related cognitive deficits. Future studies in this direction are definitely needed.

More importantly, the cross-sectional nature of this study does not allow us to give any prognostic insight into CP based on these findings, or to fully understand its translation into clinical practice potential. Longitudinal

studies are warranted to provide machine learning-based prognostic information.

CONCLUSIONS

Our findings suggest that, using psychological and physical assessments, it is possible to classify CP from HC with high reliability via (1) a pattern of psychological symptoms (identified through HADS subscales) and cognitive beliefs (identified through BMQ subscales) characterizing CP, and (2) a pattern of intact physical functions (identified through SF-36 subscales) characterizing HC. We think that our algorithm provides important and novel insights for clinical practice. Indeed, if externally validated in geographically diverse cohorts and with longitudinal information, the investigation of such psycho-physical impairments through these subscales could be prioritized, helping to better tailor early identification and intervention strategies in CP through:

– constant monitoring of the onset and the evolution of symptoms of depression and anxiety and of cognitive beliefs and disbeliefs about medicine and pharmacological treatments; and
– active promotion of physical health strategies by specifically targeting occupational, physical, and vitality impairments in CP.

This would potentially lead to improved quality of life in CP individuals and to a shorter, more transient, less burdensome pain experience.

ACKNOWLEDGEMENTS

Compliance with Ethics Guidelines. This research was conducted ethically in accordance with the World Medical Association (www.wma.net) Declaration of Helsinki of 1964 and its later amendments. All individuals gave their written informed consent before entering the study. The research protocol was approved by Bari University Hospital local ethical committee.

Data Availability. The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

REFERENCES

1. Classification of chronic pain. Descriptions of chronic pain syndromes and definitions of pain terms. Prepared by the International Association for the Study of Pain, Subcommittee on Taxonomy. Pain Suppl. 1986;3:S1–226.

2. Nadar MS, Jasem Z, Manee FS. The cognitive functions in adults with chronic pain: a comparative study. Pain Res Manag. 2016;2016:5719380.

3. Bussone GGL, Panerai AE. Pain, emotion, headache. Headacke. 2012;52:98–101.

4. Duenas M, Ojeda B, Salazar A, Mico JA, Failde I. A review of chronic pain impact on patients, their social environment and the health care system. J Pain Res. 2016;9:457–67.

5. Reid KJ, Harker J, Bala MM, Truyers C, Kellen E, Bekkering GE, Kleijnen J. Epidemiology of chronic non-cancer pain in Europe: narrative review of prevalence, pain treatments and pain impact. Curr Med Res Opin. 2011;27(2):449–62.

6. Duenas M, Salazar A, Ojeda B, Fernandez-Palacin F, Mico JA, Torres LM, Failde I. A nationwide study of chronic pain prevalence in the general spanish population: identifying clinical subgroups through cluster analysis. Pain Med. 2015;16(4):811–22.

7. Closs SJ, Staples V, Reid I, Bennett MI, Briggs M. The impact of neuropathic pain on relationships. J Adv Nurs. 2009;65(2):402–11.

8. Moriarty O, McGuire BE, Finn DP. The effect of pain on cognitive function: a review of clinical and preclinical research. Prog Neurobiol. 2011;93(3):385–404.

9. Ojeda B, Salazar A, Duenas M, Torres LM, Mico JA, Failde I. The impact of chronic pain: the perspective of patients, relatives, and caregivers. Fam Syst Health. 2014;32(4):399–407.

10. McCracken LM, Spertus IL, Janeck AS, Sinclair D, Wetzel FT. Behavioral dimensions of adjustment in persons with chronic pain: pain-related anxiety and acceptance. Pain. 1999;80(1–2):283–9.

11. Antonucci LA, Pergola G, Pigoni A, et al. A pattern of cognitive deficits stratified for genetic and environmental risk reliably classifies patients with schizophrenia from healthy control subjects. *Biol Psychiatry* 2019.

12. Dwyer DB, Falkai P, Koutsouleris N. Machine learning approaches for clinical psychology and psychiatry. Annu Rev Clin Psychol. 2018;14:91–118.

13. Kloppel S, Stonnington CM, Chu C, et al. Automatic classification of MR scans in Alzheimer's disease. Brain. 2008;131(Pt 3):681–9.

14. Zarogianni E, Moorhead TW, Lawrie SM. Towards the identification of imaging biomarkers in schizophrenia, using multivariate pattern classification at a single-subject level. NeuroImage Clin. 2013;3:279–89.

15. Antonucci LA, Penzel N, Pergola G, et al. Multivariate classification of schizophrenia and its familial risk based on load-dependent attentional control brain functional connectivity. Neuropsychopharmacology. 2020;45(4):613–21.

16. Koutsouleris N, Kahn RS, Chekroud AM, et al. Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. Lancet Psychiatry. 2016;3(10):935–46.

17. Koutsouleris N, Riecher-Rossler A, Meisenzahl EM, et al. Detecting the psychosis prodrome across high-risk populations using neuroanatomical biomarkers. Schizophr Bull. 2015;41(2):471–82.

18. Boissoneault J, Sevel L, Letzen J, Robinson M, Staud R. Biomarkers for musculoskeletal pain conditions: Use of brain imaging and machine learning. Curr Rheumatol Rep. 2017;19(1):5.

19. Lee J, Mawla I, Kim J, et al. Machine learning-based prediction of clinical pain using multimodal neuroimaging and autonomic metrics. Pain. 2019;160(3):550–60.

20. Lotsch J, Ultsch A. Machine learning in pain research. Pain. 2018;159(4):623–30.

21. Rogachov A, Cheng JC, Hemington KS, Bosma RL, Kim JA, Osborne NR, Inman RD, Davis KD. Abnormal low-frequency oscillations reflect trait-like pain ratings in chronic pain patients revealed through a machine learning approach. J Neurosci. 2018;38(33):7293–302.

22. First MB GM, Spitzer RL, Williams JBW. *Guide for the structured clinical interview for DSM-IV axis I disorders-research version*. 1996.

23. Benjamini YHY. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc: Ser B (Methodol). 1995;57(1):289–300.

24. TE Argentero P, Tibaldi G, Horne R, Clatworthy J, Munizza C. The beliefs about drug treatments. The Italian version of the BMQ (The Beliefs about Medicines Questionnaire): its validity and applicability. Epidemiologia e Psichiatria Sociale. 2011;29(01):86–92.

25. Iani L, Lauriola M, Costantini M. A confirmatory bifactor analysis of the Hospital Anxiety and Depression Scale in an Italian community sample. Health Qual Life Outcomes. 2014;12:84.

26. Apolone G, Mosconi P. The Italian SF-36 Health Survey: translation, validation and norming. J Clin Epidemiol. 1998;51(11):1025–36.

27. Basso A, Capitani E, Laiacona M. Raven's Coloured Progressive Matrices: normative values on 305 adult normal controls. Funct Neurol. 1987;2(2):189–94.

28. Ruschhaupt M, Huber W, Poustka A, Mansmann U. A compendium to ensure computational reproducibility in high-dimensional classification tasks. Stat. Appl. Genet. Mol. Biol. 2004;3:37.

29. Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. BMC Bioinform. 2006;7:91.

30. Saeys Y, Inza I, Larranaga P. A review of feature selection techniques in bioinformatics. Bioinformatics. 2007;23(19):2507–17.

31. Noble WS. What is a support vector machine? Nat. Biotechnol. 2006;24(12):1565–7.

32. Koutsouleris N, Kambeitz-Ilankovic L, Ruhrmann S, et al. Prediction models of functional outcomes for individuals in the clinical high-risk state for psychosis or with recent-onset depression: a multimodal, multisite machine learning analysis. JAMA Psychiatry. 2018;75(11):1156–72.

33. Sheng J, Liu S, Wang Y, Cui R, Zhang X. The link between depression and chronic pain: neural mechanisms in the brain. Neural Plast. 2017;2017:9724371.

34. Timmerman L, Stronks DL, Huygen FJ. The relation between patients' beliefs about pain medication, medication adherence, and treatment outcome in chronic pain patients: a prospective study. Clin J Pain. 2019;35(12):941–7.

35. Jones J, Rutledge DN, Jones KD, Matallana L, Rooks DS. Self-assessed physical function levels of women with fibromyalgia: a national survey. Womens Health Issues. 2008;18(5):406–12.

36. Fishbain DA, Cutler R, Rosomoff HL, Rosomoff RS. Chronic pain-associated depression: antecedent or consequence of chronic pain? A review. Clin J Pain. 1997;13(2):116–37.

37. de Heer EW, Gerrits MM, Beekman AT, et al. The association of depression and anxiety with pain: a study from NESDA. PLoS ONE. 2014;9(10):e106907.

38. Asmundson GJ, Katz J. Understanding the co-occurrence of anxiety disorders and chronic pain: state-of-the-art. Depress Anxiety. 2009;26(10):888–901.

39. Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain. 2000;85(3):317–32.

40. Lopez-Silva MSD, Rodriguez-Fernandez MC, Vazquez-Seijas E. Cavidol: quality of life and pain in primary care. Rev Soc Esp Dolor. 2007;14:9–19.

41. Langley P, Perez Hernandez C, Margarit Ferri C, Ruiz Hidalgo D, Lubian LM. Pain, health related quality of life and healthcare resource utilization in Spain. J Med Econ. 2011;14(5):628–38.

42. Breivik H, Eisenberg E, O'Brien T, Openminds A. The individual and societal burden of chronic pain in Europe: the case for strategic prioritisation and action to improve knowledge and availability of appropriate care. BMC Public Health. 2013;13:1229.