Research Article ISSN 2639-846X

Anesthesia & Pain Research

Personalized Pain Therapy: Artificial Intelligence (AI) Utilized to Predict Patient Response to OTC Topical Analgesics

Jeffrey Gudin[1], Seferina Mavroudi[2,3], Aigli Korfiati[3], Derek Dietze[4] and Peter Hurwitz[5*]

[1]*University of Miami School of Medicine, Miami, Florida USA.*

[2]*School of Health Rehabilitation Sciences, University of Patras, Greece.*

[3]*InSyBio Ltd, Winchester, UK.*

[4]*Metrics for Learning LLC, Queen Creek, Arizona, USA.*

[5]*Clarity Science LLC, Narragansett, Rhode Island, USA.*

[*]Correspondence:

Peter Hurwitz, Clarity Science LLC, 750 Boston Neck Road, Suite 11, Narragansett, RI 02882, Tel +1917 757 0521; Fax +1855-891-8303.

Received: 21 March 2021; **Accepted:** 25 April 2021

Citation: Gudin J, Mavroudi S, Korfiati A, et al. Personalized Pain Therapy: Artificial Intelligence (AI) Utilized to Predict Patient Response to OTC Topical Analgesics. Anesth Pain Res. 2021; 5(1): 1-11.

ABSTRACT

Purpose: *Topical analgesics have shown efficacy for patients experiencing mild and moderate pain. Due to high variability in patient demographics, clinical profile, and analgesic response, identifying the most suitable treatment for pain patients is difficult. Artificial intelligence and machine learning techniques have shown promise in individualizing treatments. This analysis reviews an interpretable machine learning method to individualize treatment. Due to adverse effects associated with many analgesics, the ability to predict treatment response has a tremendous benefit to clinicians and patients.*

Patients and methods: *Data were evaluated from 186 pain patients enrolled in an Institutional Review Board approved study (RELIEF) after use of a topical pain-relieving analgesic patch for 14 days. A novel interpretable machine learning method was developed based on a multi-objective ensemble classification/regression technique. Data was expanded to increase predictive accuracy with pre- and post-modeling techniques to raise interpretability. 85 features were identified that allowed calculation of data between testing and training groups. Data were split into training (n=152) and testing (n=34) patient sets in a stratified manner. Three basic endpoints were examined for the prediction models: total BPI Severity scores, total BPI Interference scores, and changes in the total drugs.*

Results: *Results demonstrated that the machine learning models were able to predict endpoints with extremely high accuracy, with the AUC exceeding 90% and Spearman correlation metric exceeding 0.4 for all endpoints, far exceeding the test set performance of other benchmark models. The machine learning method reduced the number of significant features from 85 to 19 and defined well characterized groups of responders and non-responders.*

Conclusion: *The machine learning model demonstrated that predictions of positive response could have been made prospectively for patients that benefited from the topical pain-relieving patch. This predictive analytic methodology can be applied to separate and larger datasets and used retrospectively to analyze whether a certain treatment might be effective in a given population.*

Keywords

Non-opioid treatment, Machine learning, Predictive analytics, Salonpas® Pain Relieving Patch, Topical analgesics.

Abbreviations: AI: Artificial Intelligence; AUC: Area Under the Curve; KNN: k-Nearest Neighbors; MSE: Mean Squared Error; ML: Machine learning; RELIEF: Relieving Pain: Evaluating Patient Quality of Life Improvement – Perceptions, Experience and Feedback After Use of A Topical Pain RELIEF Patch; OPERA: Optimizing Patient Experience and Response to Topical

Analgesics; PCA: Principal Component analysis; ROC: Receiver Operating Characteristic; SVR: Support Vector Regression.

Introduction

Chronic pain is a stressful and life-altering condition associated with various disease states [1]. It interferes with every aspect of a person's life, including general activities, walking, work, mood, enjoyment of life, relations with others, and sleep. More severe and persistent pain may also lead to a chronic pain "syndrome", described as a mental health condition going beyond pain symptoms leading to the development of conditions like depression and anxiety and being furthermore linked with mortality [2-6].

The goals of chronic pain treatments include effective pain relief, improved quality of life, and enhanced functional ability. Medications, including opioids and nonsteroidal anti-inflammatory drugs (NSAIDs), are commonly prescribed for chronic pain, but since the declaration of a public health emergency known as the "opioid epidemic" in 2017 [3] chronic pain therapy researchers have emphasized non-opioid treatments.

Topical analgesic therapies can offer pain relief devoid of the risks of abuse, misuse, and addiction [7]. The effectiveness of topical analgesics therapies for chronic pain patients was examined in our previous Institutional Review Board (IRB) -approved observational study, Optimizing Patient Experience and Response to Topical Analgesics (OPERA) [8, 9]. The results advocated for the benefits of topical analgesics therapies in reducing pain severity, improving function (reducing pain interference) and reducing patient overall analgesic drug consumption (total drugs used) during the follow-up periods.

Pain-relieving patches are a specific category of analgesics. There is a distinction between transdermal patches- where the medication is absorbed into the systemic circulation (e.g. fentanyl, nicotine), and topical patches, such as lidocaine and methyl salicylate. The Salonpas® Pain Relieving Patch (Hisamitsu Pharmaceutical Company, Inc, Japan), is an over the counter (OTC) analgesic topical pain patch that includes menthol, camphor, and methyl salicylate being recently shown to reduce the severity and interference of pain in chronic pain patients [10].

Our observational outcomes [9,10] have shown that even though topical analgesics are beneficial for the majority of the patients studied, the effectiveness varies, and a subset of patients do not benefit at all. Thus, a mechanism to predict analgesic responders would be of great value by allowing clinicians to identify the most suitable analgesic therapy for each chronic pain patient. Many research studies [11-13] have attempted to individualize pain treatments via pharmacogenetics and other patient phenotypes or characteristics; as of now, there remains no validated way for clinicians to predict individual patient responses in order to select the most effective among different treatment plans and to make informed critical clinical decisions. A machine learning analytic approach has been utilized previously [14] to predict chronic pain patient response to topical analgesic treatment. The results suggested that this machine learning model could have predicted in advance at least 10% of patients who (would have) failed treatment with the studied therapy.

One of the main limitations of machine learning applications in translational research is their complexity and the fact that no clear conclusions can be made for which features are the most defining ones for each prediction. Recently, a new category of techniques has been suggested, Explainable Artificial Intelligence (XAI) methods [15-17], whose need became obvious by the documentation of decision errors of AI systems attributed to biases embedded in the training data [18]. Besides detecting vulnerable points and helping to improve prediction models, explainability is also essential because it produces knowledge about domain relationships contained in the data that may help to unravel the underlying explanatory factors of the data.

In the present paper, we introduce a new interpretable machine learning pipeline, enrich it with pre- and post-modelling explainability modules, and apply it to individualize the therapy of chronic pain treatment by predicting the benefits of the Salonpas® Pain Relieving Patch treatment.

Material and methods

Data

In the present analysis, data from the RELIEF study [10] were used to individualize the topical analgesics therapy by training and testing prediction models to evaluate which chronic pain patients would benefit from the Salonpas® Pain Relieving Patch treatment. Three different endpoints were examined in the present study including the changes in 1) total severity, 2) total interference and 3) total drugs before and after the treatment with Salonpas® Pain Relieving Patch from day 0 to day 14. Total severity and total interference were estimated using the Brief Pain Inventory (BPI) scale [19]. The RELIEF study was performed in full accordance with the rules of the Health Insurance Portability and Accountability Act of 1996 (HIPAA) and the principles of the declaration of Helsinki and the international council of Harmonisation/Good Clinical Practice (GCP). The study protocol was approved by the IntegReview institutional review board.

The RELIEF dataset was composed of data before and after treatment for 186 chronic pain patients. 89% of participants improved their total severity and total interference status and 42% reduced the total number of analgesic drugs used after study patch treatment employed. The dataset was split to training and test sets with training set having 152 patients and test set 34 patients maintaining the same percentages of responders and non-responders to treatment. All patients replied to 41 questions before and 14-days after the treatment. These questions were formulated as features by encoding categorical variables to features using the FeatureHasher method of the Scikit-learn package [20] to allow for their integration with number features, ending up with a list of 85 features. Data were scaled to zero means and standard deviation of 1 and missing values were imputed with KNN-imputation method using k=5. 10-fold stratified cross validation was used on

the training set to maintain the same percentage of responders and non-responders in the different folds.

Machine Learning Method

The three prediction problems were treated as regression problems since the outcomes of this study can be real numbers.

The machine learning method utilized was applied previously in another topical analgesics population study [14]. We attempted to further improve its ability to train and test accurate regression models with imbalanced datasets minimizing the number of selected features and improving interpretability of the models. In particular, an ensemble dimensionality reduction technique employing a multi-objective heuristic optimization algorithm [21] allowed us to a) identify the optimal feature subset to be used as input to the classifiers and to select b) the most appropriate regression model among Support Vector Regression [22] and Random Forest Methods [23], c) its optimal parameters and d) the optimal order of the classification models in classifiers chain. Classifiers chains method [24] was used to take into consideration interactions between the examined endpoints.

The heuristic optimization framework is a pareto-optimization technique since its selection process is driven by organizing solutions to non-dominated fronts and assigning close fitness values to solutions belonging to the same front.

The iterative process of the optimization framework begins by initializing a set of solutions. Each solution consists of i) a value indicating which of the two alternative Support Vector Regression kernel types and the Random Forest models will be used (a value in [0,1) indicates the selection of Radial Basis Function Kernel SVR, a value in [1,2) indicates the selection of linear Kernel SVR and a value in [2,3) indicates using random forests), ii) 85 values for deciding if a feature will be used as input (values greater than 0.5 force its use), iii) three values for optimizing the gamma parameter of Radial Basis Functions Kernel, the regularization parameter C of SVR models and the number of Random Trees in Random Forests and iv) one variable for the selection of which of the 6 potential rankings of the predictors will be used in the classifiers chain. The first population of solutions is generated by randomizing values considering normal distribution of each variable.

The optimization goals that were formulated as Fitness Functions were following:

- **Fitness Function 1:** Classification accuracy
- **Fitness Function 2:** Area under the Receiver Operating Characteristic (ROC) curve
- **Fitness Function 3**: Spearman correlation in between predicted and real outputs
- **Fitness Function 4:** minimalization of support vectors or random tree)
- **Fitness Function 5:** 1/(1+number of selected features)

The outputs of the prediction models were binarized to class 1 if the value is bigger than or equal to 0, and to class -1 otherwise to calculate the classification metrics in the aforementioned fitness functions.

The utilized fitness functions aim to minimize the number of selected features, maximize regression and classification performances and minimize the complexity of the classifier.

After the evaluation of the population, the Pareto fronts of non-dominated solutions are calculated and solutions are assigned a fitness value based on their pareto front. The Roulette Wheel Selection method is applied to generate a new population of solutions which are then differentiated using the Genetic Algorithms two-point crossover and Gaussian Mutation operators. The new population is evaluated again and this iterative process continues until it converges (solutions become close enough for a number of iterations) or reaches the maximum number of generations (Figure 1).

In order to further interpret the revealed prediction models, the selected features were further analyzed using Spearman Correlation and hierarchical clustering as well as Principal Component Analysis and unsupervised K-Prototypes clustering of the patients combined with functional enrichment analysis. The k in k-prototypes Clustering is based on the k-prototypes algorithm [25] because we have both numerical and categorical inputs. Calinski-Harabasz score [26] was used to calculate the optimal number of clusters experimenting for number of clusters in between 2 and 20.

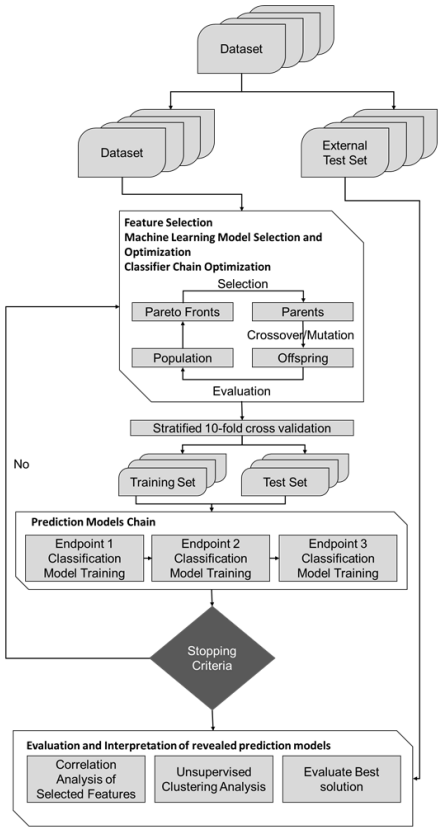


Figure 1: Flowchart of the proposed machine learning model.

To compare our method against, SVR and Random Forests models were trained and evaluated using WEKA software (we used the default parameters suggested in WEKA documentation) [27] using the same training and test sets and the same cross validation strategy as in the proposed machine learning method, while also optimizing their parameters (C and gamma for SVR, and number of trees for Random Forests) with grid search.

Results

PCA and Clustering Analysis

Principal component analysis (PCA) and k-Prototypes unsupervised clustering was conducted to explore whether the RELIEF data can be separated in clusters that can discriminate between responders and non-responders (Figure 2). The best clustering based on Calinski Charabazs metric was conducted when two clusters were used. By examining the projections of the endpoint changes it is easily observable that the two revealed clusters have some potential into discriminating responders and non-responders based on the BPI Severity and BPI Interference changes but not on the Total drugs change. Moreover, it seems that there does not exist a single hyperplane differentiating between responders and non-responders for none of the examined endpoints and thus non-linear methods are required to optimize classification.

Predictive Analytics and Comparative Results

The proposed machine learning model was applied in the examined dataset using as parameters: population size of 100 and maximum number of generations of 200. Since the proposed method is a heuristic approach, it was applied 10 times in the proposed dataset and Figure 3 presents its average performance in predicting the 3 examined endpoints. Figure 3 also presents the performance of state-of-the-art regression models, SVR and Random Forests,

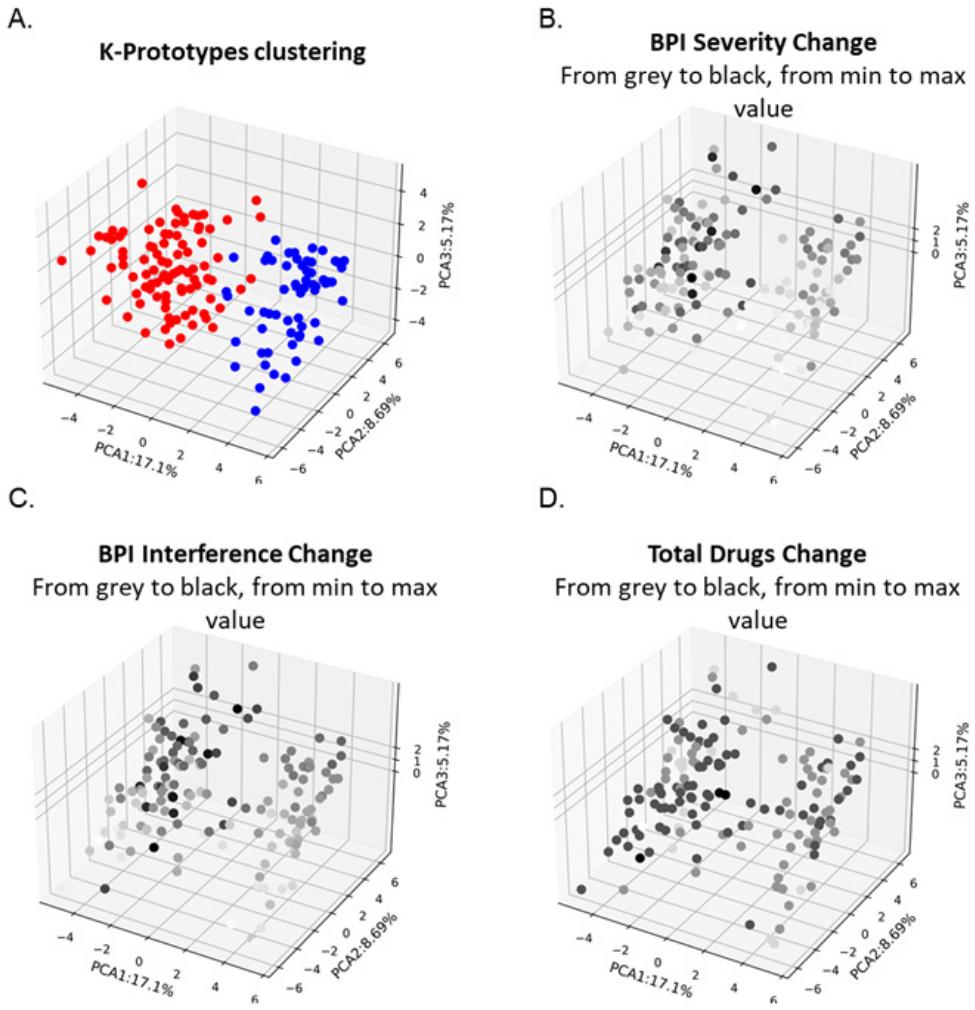


Figure 2: A. K-Prototypes clustering of RELIEF Dataset. Red and Blue nodes depict the revealed two clusters using the three more important Principal Components and the percentage of the explained variability of each one of these PCAs is depicted in the axis labels. B. 3D representation of the PCAs of the RELIEF Dataset projecting on the samples the BPI Severity Change before and after the treatment. Grey to black color scale was used to depict values from min to max respectively. C. 3D representation of the PCAs of the RELIEF Dataset projecting on the samples the BPI Interference Change before and after the treatment. Grey to black color scale was used to depict values from min to max respectively. D. 3D representation of the PCAs of the RELIEF Dataset projecting on the samples the Total Drugs Change before and after the treatment. Grey to black color scale was used to depict values from min to max respectively. *PCA- Principal Component Analysis; BPI- Brief Pain Inventory.

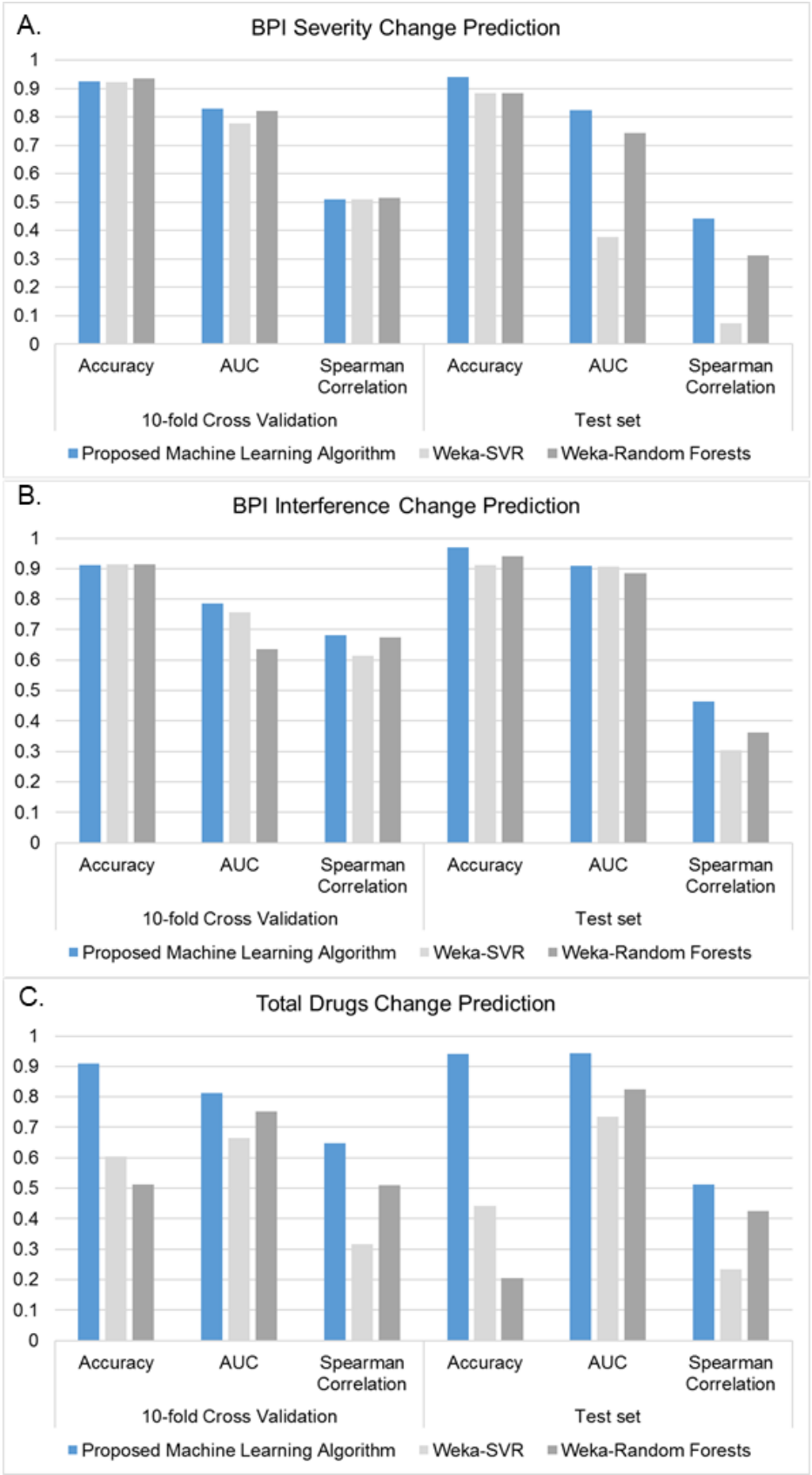


Figure 3: Comparative results of trained machine learning models in predicting A. Total Severity Change, B. Total Interference Change and C. Total Drugs Change. Evaluation metrics have been calculated in Training set using 10-fold cross validation and in Test set using the external test set. *BPI-Brief Pain Inventory.

implemented through Weka software and with their parameters being optimized using grid search and cross validation in the training dataset.

Experimental results show the superiority of the proposed machine learning method in all examined metrics with the increase being even more pronounce in the external test set. It is noteworthy, that the proposed model achieved high Spearman correlation values (higher than 0.4) of the predicted versus the real output surpassing the other methods for all endpoints. It is noteworthy that the improved performance of the proposed method came with a significant reduction in the utilized features since it used only 19 of the 85 features of the dataset.

The best predictor that was found using the proposed machine learning method was a Random Forest of 47 random trees that uses 19 features while the order of the predictors in the classifiers chain starts from the predictor of BPI Interference change, followed by the predictor of BPI Severity change and the predictor of Total drugs change respectively.

The final trained predictive model is accessible as trained Sklearn Random Forest model in https://www.insybio.com/pain_research/

Interpreting Prediction Models

Table 1 presents the selected features and their category. 19 features were selected originating from 16 questions from the Primary complaint/diagnosis and location, the brief pain inventory and the current medications categories.

Spearman correlation and hierarchical clustering analysis was conducted to further explore why these features were important and interpret the trained machine learning models. In particular the selected features were cross correlated and then hierarchical clustering was applied revealing that these features are organized in three uniform clusters. The first cluster is negatively correlating with the endpoint's change, the second is positively correlating with the endpoints' change, while the third cluster is a mixed one.

As another step of for the interpretation of the revealed features set and model, unsupervised k-prototypes clustering was applied to applied to the RELIEF dataset similarly to the analysis that was done and presented in Figure 2 but using only the 19 selected features (Figure 4). Calinski-Harabazs metric was used to identify the optimal number of clusters examining clusterings with 2 to 20 clusters. The best clustering was achieved for 7 clusters and it is presented in Figure 5. In order to better understand the association of the features with the outcomes we performed enrichment analysis of the revealed clusters with the binarized endpoints variables and annotated the clusters with the examined features using the centers of the revealed clusters. This analysis showed that 5 out of 7 clusters of participants decreased their BPI scores after treatment but this happened only for 2 out for 7 clusters for the total number of medications change. Moreover, a cluster of super-responders (Cluster 1) who reduced both BPI scores and total number of drugs was revealed consisting of 16.2% of the participants. These super-responders were Voltaren® users, taking high number of anti-inflammatories, having problems in sleep and doing no or very low physical exercise. This group is predicted to get the most benefits from initiating a Salonpas® Pain Relieving Patch -based pain treatment. On the other hand, there is a small clear cluster of non-responders (Cluster 6) who did not have an office staff verification of pain complaint diagnosis, are already using alternative medications and do not have knee pain.

Examining the Applicability of the Individualized Prediction Models in Other Cohorts

From a previous study of our authoring group [14] a machine learning model was used to identify responders of topical analgesics therapies using the OPERA study dataset, which is a dataset of 631 chronic pain patients also treated with a topical analgesic [8,9]. In this study approximately 10% of the participants had been predicted not to be suitable for the examined topical analgesics therapies.

In the context of the present study, an interpretable machine learning model was developed to predict the specific response of chronic pain patients in Salonpas® Pain Relieving Patch -based treatment for chronic pain. In order to explore how this model can be applied in another cohort we used the trained model for Salonpas® treatment outcome prediction to the data of OPERA study. Since, two of the features of the Salonpas® treatment outcome prediction model (BQ2, Bq6.3) were not measured in the OPERA study, k-NN imputation was used for them after merging the RELIEF and OPERA datasets using their common features. Prediction analysis showed that out of the 10% of non-responders from the topical analgesic therapies in the OPERA study, 73.6% of them would have presented reduced BPI scores if they were treated with Salonpas® Pain Relieving Patch. Moreover, 62.3% of the responders of topical analgesics treatments in OPERA study would have further decreased the overall response measures as the average of the reductions in BPI scores and Total number of drugs administered to them.

Discussion

Chronic pain affects tens of millions of people worldwide, affecting function and quality of life. In light of the opioid crisis, emphasis has been placed on the development of non-opioid therapies, including topical analgesics patches, and exploring their potential to alleviate chronic pain. However, these therapies are not beneficial for all chronic pain patients and practitioners would benefit from a predictive, personalized medicine approach to determine which therapy should be administered to each patient based on their clinical and demographic profile. In the present analysis, we attempted to evaluate the use of explainable machine learning into building accurate outcome prediction models for a topical analgesic patch for chronic pain patients.

PCA analysis using the features collected in the RELIEF study demonstrated that these features have potential to discriminate the between responder and non-responder pain patients for the

Table 1: Final list of selected features from machine learning model.

A/A	Feature_id	Question	Category
1	Sum_Arthritis	Arthritis	Primary Complaint/ Diagnosis and Location
2	BQ2	Office Staff Verification of Primary Pain Complaint Diagnosis	Primary Complaint/ Diagnosis and Location
3	BQ6.3	Number of Weekly Heavy Physical Activity for 30 minutes or more	Primary Complaint/ Diagnosis and Location
4	BBPI1	Pain other than everyday kinds of pain	Brief Pain Inventory
5	BPI2.1OPEN_2	Areas of Pain (Hashing feature 2)	Brief Pain Inventory
6	BPI2.1OPEN_3	Areas of Pain (Hashing feature 3)	Brief Pain Inventory
7	BPI2.1OPEN_4	Areas of Pain (Hashing feature 4)	Brief Pain Inventory
8	BPI2.1OPEN_5	Areas of Pain (Hashing feature 5)	Brief Pain Inventory
9	BBPI3	Worst Pain	Brief Pain Inventory
10	BBPI4	Least Pain	Brief Pain Inventory
11	BBPI5	Average Pain	Brief Pain Inventory
12	BBPI9.5	Relations with other people	Brief Pain Inventory
13	BBPI9.6	Sleep	Brief Pain Inventory
14	BBPIInterference	BBPI Interference in Baseline	Brief Pain Inventory
15	Bmed1.5	Other medications	Current Medications
16	BPmed2.15	Voltaren	Current Medications
17	BNumberOTC	Number of Over the Counter Medications	Current Medications
18	BNumberAntiinflammatoryRx	Number of Anti-inflammatory Medications	Current Medications
19	BTotalNumberMedications	Total Number of Medications	Current Medications

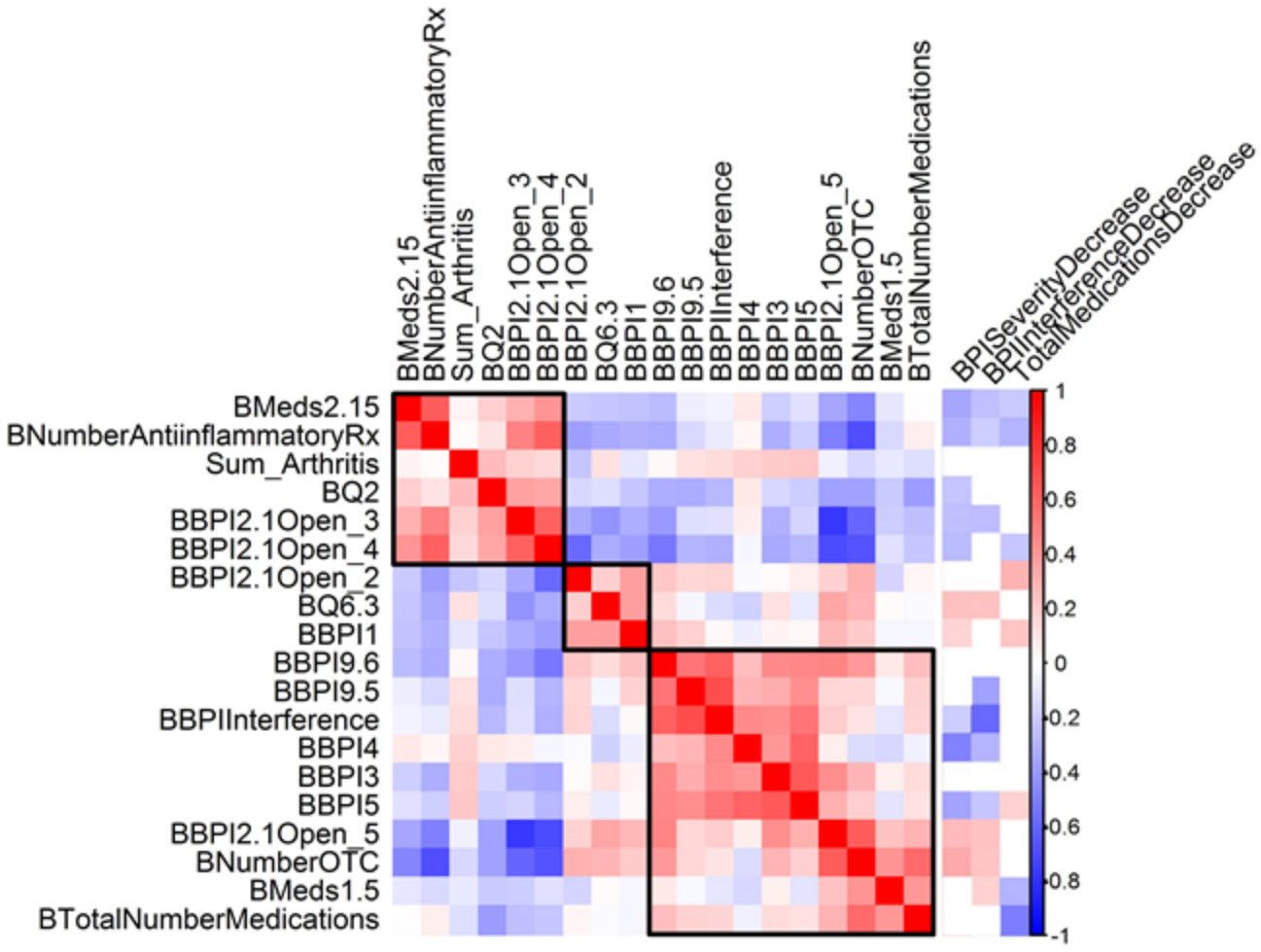
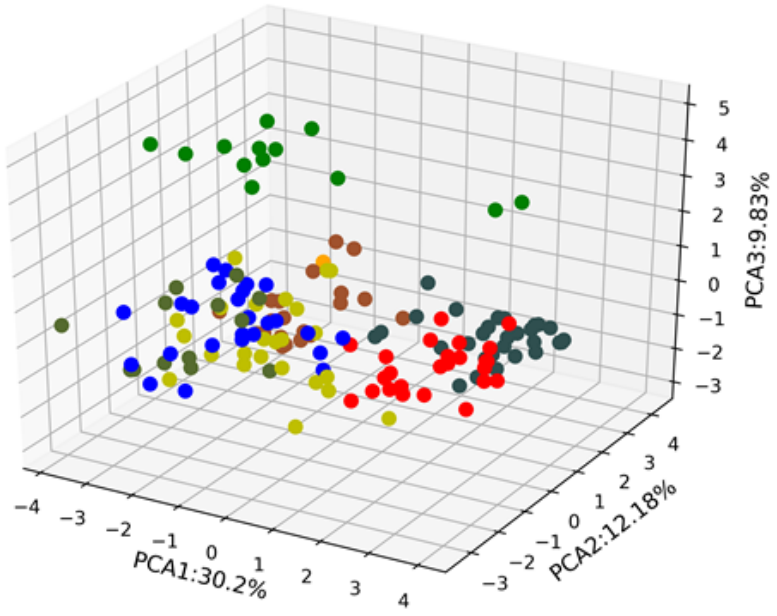


Figure 4: Spearman correlation in between selected features from proposed machine learning model. Rectangles depict the revealed clusters of features using hierarchical clustering while the Spearman correlation of each feature with the examined endpoints is also provided. Insignificant correlations are denoted with an empty cell in the heatmap.



Cluster	Number of Patients	BPI Severity & Interference Changes	Total Medications Change	Clinical, Demographics Characteristics based on the clusters centers
Cluster 1	25	Decreased	Decreased	Voltaren users, high number of anti-inflammatories, problems with sleep, low number of over the counter medications, low or no heavy activity
Cluster 2	25	Decreased	Mixed	Voltaren users, high number of anti-inflammatories, arthritis patients
Cluster 3	24	Decreased	Not Decreased	Voltaren users, high number of anti-inflammatories, office staff verification of primary pain complaint diagnosis
Cluster 4	13	Not Decreased	Decreased	Low initial BPI interference score, Absence of office staff verification of primary pain complaint diagnosis, using other medications, no foot pain, no knee pain
Cluster 5	20	Decreased	Mixed	Office staff verification of primary pain complaint diagnosis, no arthritis patients, no anti-inflammatories, low or no heavy physical activity
Cluster 6	13	Mixed	Not Decreased	Absence of office staff verification of primary pain complaint diagnosis, high number of over the counter medications, no knee pain
Cluster 7	31	Decreased	Mixed	High number of anti-inflammatories, office stafo verification of primary pain complaint, not taking other medications

Figure 5: 3D PCA representation of the participants of RELIEF Patients using the selected features from the proposed machine learning method. Clustering was conducted using k-prototypes method and the revealed clusters characteristics were presented using the centers of each cluster and hypergeometric distribution enrichment analysis was used to annotate clusters based on the examined endpoints.

Salonpas® Pain Relieving Patch, but there does not exist a simple linear model that could classify responders and non-responders with adequate accuracy. Thus, more elaborate machine learning methods were required for this task. However, machine learning methods have the limitations of being mostly viewed as black boxes by clinicians and non-domain experts limiting thus their translational applications and significantly raising the adoption obstacles. For this reason, the present study introduced a new ensemble multi-objective optimization regression/classification machine learning model which uses contemporary pre- and post modelling explainability techniques to raise the interpretability of the final models. The applied method has been proven to be efficient in reducing the dimensionality of the prediction problem by selecting the most suitable feature subset, reducing thus the complexity of the overall model and raising its generalization properties. This pipeline was integrated with visualization, correlation, unsupervised clustering and enrichment analysis techniques to further explore the relationship of the selected inputs with the examined endpoints of the study shedding light into the trained prediction models.

Experimental results demonstrated that the proposed method significantly surpassed the performance of contemporary benchmark machine learning models in all the examined metrics. This improvement in the performance is attributed to the advanced feature selection mechanism and to the selection of the most suitable classification model and the optimization of its parameters for the selected feature subset. By forcing the trained model to be as simple as possible minimizing the selected features and the number of random trees or support vectors for RF and SVR methods respectively, the proposed method achieved significantly higher generalization properties in the independent test set compared to the benchmark models.

Post-modelling analysis revealed the relationships of the selected features between them and against the endpoints of the study as well as well-defined and characterized clusters of responders and possible non-responders of the Salonpas® Pain Relieving Patch chronic pain treatment in the group of patients studied. The Salonpas® Pain Relieving Patch seems to provide maximal benefits in patients who take high number of anti-inflammatory drugs, use Voltaren®, have pain in knee or foot, do little or no heavy physical exercise and have problems with their sleep. On the other hand, they do not seem to be effective for patients who do not have an authorized diagnosis of chronic pain and already use alternative over the counter medications. Single feature correlation analysis demonstrated that the higher the initial total interference score the bigger the margin for improvement using RELIEF study data while the higher the number of medications will lead to a bigger reduction of medications using Salonpas® Pain Relieving Patch but without reducing substantially the BPI scores.

To further evaluate the performance of the trained outcome prediction models for the Salonpas® Pain Relieving Patch into individualizing the therapy of chronic pain patients we applied the trained models to predict responders and non-responders of Salonpas® Pain Relieving Patch in the participants of OPERA study. This analysis demonstrated that the majority of the participants in OPERA who were proven not to be benefited from the examined topical analgesics treatments in this study would have been benefited from using Salonpas® Pain Relieving Patch. Moreover, 62.3% of the responders of topical analgesics treatments in OPERA study would have further decreased the overall response measures, such as the average of the reductions in BPI scores and the Total number of drugs administered to them.

The present study presents several limitations considering the availability of data and the limitations of the analysis techniques. The RELIEF study was an IRB-approved observational study and limited information was available about the specific diseases that were the cause of the chronic pain condition of the participants. The lack of this information not only limited the predictive accuracy of the models but did not allow us to explore the systemic effect of the topical analgesic treatments on the cause of the disease if any such exists. Moreover, the absence of in-depth follow-up of the patients for a longer period of time (e.g., 1-3 years) further restricted the assessment of the systemic and permanent effect of topical analgesics treatments. The assessment of the participants of OPERA study with the trained model for individualizing the Salonpas® Pain Relieving Patch treatment of chronic pain patients was very useful to evaluate the usability of this model in an independent study and to explore combining more than one machine learning models for individualizing chronic pain treatment. However, the two studies, RELIEF and OPERA, had different follow-up timings (14 days, and 3-6 months, respectively) and thus a unified study is required to evaluate the performance of the combination of the two prediction models. Finally, from a methods point of view, the proposed machine learning model could be expanded with additional methods from the explainable AI domain [28-31].

Conclusions

Topical analgesics can be of benefit to many, but not all pain patients as demonstrated by the present and previous studies. AI, machine learning, and predictive analytic models can be successfully incorporated into medical decision making and represent a novel precision medicine approach to allow healthcare practitioners to select only the treatments that have a high probability of success. However, their black-box nature and limited generalizability properties has limited their application in clinical trial. In the present paper, an explainable machine learning technique was introduced and used to predict patients' response to OTC topical analgesics overcoming the aforementioned limitations of machine learning techniques. The proposed method, among others, revealed a group of super responders with well-defined clinical characteristics who are predicted to get the most benefits from the Salonpas® Pain Relieving Patch -based pain treatment.

Explainable machine learning models may enable clinicians to select viable nonopioid therapies with more confidence identifying the characteristics of responders and provide an unbiased method to precisely identify responders of non-opioid treatments.

Combining questionnaire and clinical data with omics data, such as pharmacogenomics, should advance the science and safety of pain management in the future by identifying more accurate and informative models for the prediction of the response of opioid and non-opioid chronic pain treatments.

Disclosure

This IRB-approved Study was administered by Clarity Science LLC and was funded through a Grant from Hisamitsu America, the distributor of the Salonpas® Pain Relieving Patch.

A. Korfiati and S. Mavroudi are the named inventors of the provisional patent (Theofilatos, K., Alexakos, C., Korfiati, A., Dimitrakopoulos, C., & Mavroudi, S. (2018). U.S. Patent Application No. 15/837,407.) submitted to the US Patent Office by InSyBio Ltd which includes the description of the computational framework for predictive biomarkers and building predictive models for diagnosis, prognosis, and treatment. Peter Hurwitz is President of Clarity Science LLC. The remaining authors and Investigators did not receive any payment from Hisamitsu America for their participation in this study or the development of the manuscript and report no further conflicts of interest in this work.

References

1. Nielsen LM, Olesen AE, Branford R, et al. Association between human pain-related genotypes and variability in opioid analgesia: an updated review. Pain Pract. 2015; 15: 580-594.
2. Torrance N, Elliott AM, Lee AJ, et al. Severe chronic pain is associated with increased 10-year mortality. a cohort record linkage study. Eur J Pain. 2010; 14: 380-386. Fine PG. Long-term consequences of chronic pain: mounting evidence for pain as a neurological disease and parallels with other chronic disease states. Pain Med. 2011; 12: 996-1004.
3. Grimby-Ekman A, Gerdle B, Björk J, et al. Comorbidities, intensity, frequency and duration of pain, daily functioning and health care seeking in local, regional, and widespread pain - a descriptive population-based survey (SwePain). BMC Musculoskelet Disord. 2015; 16: 165.
4. Reid KJ, Harker J, Bala MM, et al. Epidemiology of chronic non-cancer pain in Europe: narrative review of prevalence, pain treatments and pain impact. Curr Med Res Opin. 2011; 27: 449-462.
5. de Sola H, Salazar A, Dueñas M, et al. Nationwide cross-sectional study of the impact of chronic pain on an individual's employment: relationship with the family and the social support. BMJ Open. 2016; 6: 12246.
6. Gerdle B, Björk J, Cöster L, et al. Prevalence of widespread pain and associations with work status: a population study. BMC Musculoskelet Disord. 2008; 9: 102.
7. Casale R, Symeonidou Z, Bartolo M. Topical treatments for localized neuropathic pain. Curr Pain Headache Rep. 2017; 21: 15. Flores MP, Castro AP, Nascimento Jdos S. Topical analgesics. Rev Bras Anestesiol. 2012; 62: 244-252.
8. Gudin JA, Brennan MJ, Harris ED, et al. Changes in pain and concurrent pain medication use following compounded topical analgesic treatment for chronic pain: 3- and 6-month follow-up results from the prospective, observational Optimizing Patient Experience and Response to Topical Analgesics study. J Pain Res. 2017; 10: 2341-2354.
9. Gudin JA, Brennan MJ, Harris ED, et al. Reduction of opioid use and improvement in chronic pain in opioid-experienced patients after topical analgesic treatment: an exploratory analysis. Postgrad Med. 2018; 130: 42-51.
10. Gudin JA, Dietze DT, Hurwitz PL. Improvement of Pain and Function After Use of a Topical Pain Relieving Patch: Results of the RELIEF Study. Journal of Pain Research. 2020; 13: 1557-1568.
11. Forstenpointner J, Otto J, Baron R. Individualized neuropathic pain therapy based on phenotyping: are we there yet?. Pain. 2018; 159: 569-575.
12. Gwynne-Jones DP, Gray AR, Hutton LR, et al. Outcomes and factors influencing response to an individualized multidisciplinary chronic disease management program for hip and knee osteoarthritis. The Journal of arthroplasty. 2018; 33: 2780-2786.
13. Emery MA, Akil H. Endogenous Opioids at the Intersection of Opioid Addiction, Pain, and Depression: The Search for a Precision Medicine Approach. Annual Review of Neuroscience. 2020; 43: 355-374.
14. Gudin J, Mavroudi S, Korfiati A, et al. Reducing Opioid Prescriptions by Identifying Responders on Topical Analgesic Treatment Using an Individualized Medicine and Predictive Analytics Approach. Journal of Pain Research. 2020; 13: 1255-1266.
15. Tjoa E, Guan C. A survey on explainable artificial intelligence (xai): Towards medical xai. arXiv preprint arXiv:1907.07374. Xie, Y., Gao, G., & Chen, X. A. (2019). Outlining the Design Space of Explainable Intelligent Systems for Medical Diagnosis. arXiv preprint arXiv:1902.06019.
16. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nature medicine. 2019; 25: 44-56.
17. Caruana R, Lou Y, Gehrke J, et al. Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. 2015.
18. Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning. arXiv. 2017.
19. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. Annals Academy of Medicine Singapore. 1994; 23: 129-138.
20. Pedregosa F, Varoquaux G, Gramfort A, et al. Machine learning in Python. the Journal of machine Learning research. 2011; 12: 2825-2830.
21. Corthésy J, Theofilatos K, Mavroudi S, et al. An adaptive

pipeline to maximize isobaric tagging data in large-scale MS-based proteomics. Journal of proteome research. 2018; 17: 2165-2173.

22. Smola AJ, Schölkopf B. A tutorial on support vector regression. Stat Comput. 2004; 14: 199-222.
23. Breiman L. Random forests. Machine learning. 2001; 45: 5-32.
24. Read J, Pfahringer B, Holmes G, et al. Classifier chains for multi-label classification. Machine learning. 2011; 85: 333.
25. Ji J, Bai T, Zhou C, et al. An improved k-prototypes clustering algorithm for mixed numeric and categorical data. Neurocomputing. 2013; 120: 590-596.
26. Caliński T, Harabasz J. A dendrite method for cluster analysis. Commun Stat Theory Methods. 1974; 3: 1-27.
27. Frank E, Hall M, Holmes G, et al. Weka-a machine learning workbench for data mining. In: Data mining and knowledge discovery handbook. Boston MA: Springer. 2009; 1269-1277.
28. Melis DA, Jaakkola T. Towards robust interpretability with self-explaining neural networks. In Advances in Neural Information Processing Systems. 2018.
29. Al-Shedivat M, Dubey A, Xing EP. Contextual explanation networks. 2017.
30. Montavon G, Binder A, Lapuschkin S, et al. Layer-wise relevance propagation: an overview. In Explainable AI: Interpreting, Explaining and Visualizing Deep Learning. Springer Cham. 2019.
31. Higgins I, Amos D, Pfau D, et al. Towards a definition of disentangled representations. 2018.