Machine learning approach to predict medication overuse in migraine patients

Patrizia Ferroni, Fabio M. Zanzotto, Noemi Scarpatò, Antonella Spila, Luisa Fofi, Gabriella Egeo, Alessandro Rullo, Raffaele Palmirotta, Piero Barbanti, Fiorella Guadagni, Gabriella Egeo, Antonella Spila, Luisa Fofi, Patrizia Ferroni, Noemi Scarpatò, Antonella Spila, Luisa Fofi, Gabriella Egeo, Alessandro Rullo, Raffaele Palmirotta, Piero Barbanti, Fiorella Guadagni

A R T I C L E   I N F O

Article history:
Received 28 October 2019
Received in revised form 19 May 2020
Accepted 5 June 2020
Available online 12 June 2020

Keywords:
Migraine
Medication overuse
Artificial intelligence
Machine learning
Decision support systems

A B S T R A C T

Machine learning (ML) is largely used to develop automatic predictors in migraine classification but automatic predictors for medication overuse (MO) in migraine are still in their infancy. Thus, to understand the benefits of ML in MO prediction, we explored an automated predictor to estimate MO risk in migraine. To achieve this objective, a study was designed to analyze the performance of a customized ML-based decision support system that combines support vector machines and Random Optimization (RO-MO). We used RO-MO to extract prognostic information from demographic, clinical and biochemical data. Using a dataset of 777 consecutive migraine patients we derived a set of predictors with discriminatory power for MO higher than that observed for baseline SVM. The best four were incorporated into the final RO-MO decision support system and risk evaluation on a five-level stratification was performed. ROC analysis resulted in a c-statistic of 0.83 with a sensitivity and specificity of 0.69 and 0.87, respectively, and an accuracy of 0.87 when MO was predicted by at least three RO-MO models. Logistic regression analysis confirmed that the derived RO-MO system could effectively predict MO with ORs of 5.7 and 21.0 for patients classified as probably (3 predictors positive), or definitely at risk of MO (4 predictors positive), respectively. In conclusion, a combination of ML and RO – taking into consideration clinical/biochemical features, drug exposure and lifestyle – might represent a valuable approach to MO prediction in migraine and holds the potential for improving model precision through weighting the relative importance of attributes.

© 2020 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Migraine is a primary form of headache characterized by recurrent episodes of debilitating headache, sometimes preceded by transient neurological symptoms named aura. Its pathophysiology recognizes a unique mixture of bio-psycho-social aspects, which may all trigger the attack in susceptible individuals, unveiling a biological predisposition of a dysexcitable brain to convert non-painful stimulation into headache pain. This ultimately leads to impressive disability, significant productivity loss, huge economic burden and healthcare resource use.

In the last two decades, an unprecedented number of studies have contributed to substantial advances in abortive and preventive therapy, opening new and promising scenarios in the field of precision medicine. Nonetheless, migraine remains a major clinical...
challenge, as it is often under-diagnosed [1]. Current validated
diagnostic criteria, in fact, distinguish migraine according to the
attack frequency (episodic or chronic) or to the presence/absence
of aura [2], but do not disentangle the different endophenotypes
of this highly heterogeneous headache disorder [3]. Furthermore,
preventive migraine medications are often underutilized, leaving
patients at-risk of medication overuse (MO), disease progression,
higher disability and increased healthcare costs [4]. MO affects
approximately 15% of migraine patients [5] and is the most rele-
vant risk factor for migraine chronicization [6] as well as the devel-
opment of a secondary headache disorder known as medication-
overuse headache (MOH) [2,7,8]. Daily or weekly analgesic use,
in fact, disproportionately increases the risk for chronic migraine
(RR = 13.3; 95% CI: 9.3–19.1) as documented by a large prospective
population survey [9]. When considering a specialized headache
center setting, the odds for developing chronic migraine 1 year
later in patients with episodic migraine with MO is 19.4 times
(95% CI: 8.7–43.2) higher than in those without [10]. The risk of
chronic migraine progression is further incremented when overus-
ing involves acute medications containing barbiturates (OR = 2.1;
95% CI: 1.3–3.1) or opioids (OR = 1.2; 95% CI: 1.4–2.2), compared
with acetaminophen, as hinted by the American Migraine Preva-
ience and Prevention (AMPP) study [11]. This is also true in adoles-
cents, in whom medication overuse is a predictor of migraine
clinical evolution (HR = 2.5; 95% CI: 1.1–5.5) [12]. Once estab-
lished, MOH is the epitome of hard to treat neurological disorders
due to its disappointing response to treatment and frequent relapse.

Thus, the possibility of predicting MO, or estimating the efficacy
of a given therapy, is a compelling challenge in the clinical man-
agement of migraine to set out prevention programs for patients
prone to frequent headache.

Artificial intelligence (AI) with machine learning (ML) has
shown great potential in building automatic predictors in the field
of migraine [13–22] but detectors for MO are still in their infancy
[20]. In fact, many AI models have been applied for nearly 10 years
to implement medical decision support systems for the diagnosis
of migraine [13–20] and for predicting migraine treatment out-
comes [21,22]. For example, artificial neural networks (ANNs), arti-
ficial immune systems (AIS) and support vector machines (SVM)
have shown encouraging results. Some systems were built on clinical
data derived from patients’ medical records [13–17], while others
analyzed data from resting-state functional magnetic reso-
nance imaging (rs-fMRI) [18–20]. All these studies demon-
strated that intelligent systems represent a promising approach
for migraine classification, holding potential to revolutionize con-
ventional models of diagnostic deliverance in a contest of person-
alized medicine. However, To the best of our knowledge, there is
only one study studying the potential of AI in predicting MO
[20]. In their study, Garcia-Chimeno et al. reported the possibility
to distinguish between patients with episodic or chronic migraine
with MO via feature selection techniques and machine learning
[20]. Using SVM, Boosting and Naive Bayes they obtained an over-
all classification with over 93% accuracy. However, the study had
a big major limitation: due to the small number of participants
(n = 18 for chronic migraine with MO), the testing was undertaken
in the same dataset used for training and no holdout test set was
used [20]. Moreover, the specific aim was not to predict MO risk,
but to correctly classify between different migraine phenotypes
[20]. Thus, there is still an unmet need to develop classifications
models that embody the newest AI technologies and can be used
to predict MO in individual migraine patients for a personalized
patient care.

In this study, we present our approach RO-MO to use machine
learning along with Random Optimization (RO) for predicting
Medication Overuse (MO) in migraine patients. Our approach in
this study solves the problems of previous studies [20] as testing
has been carried out in a held-out set and RO-MO outperforms
SVM as used in [20]. RO-MO derives from our previous studies.
In a perspective of predictive medicine and tailored therapy, we
have recently demonstrated that an approach based on AI method-
ologies holds the potential to devise decision support systems
(DSSs) that can be adapted to different medical problems [23–
26]. Our model based on Multiple Kernel Learning (MKL) [27],
combines SVM [28] algorithm and Random Optimization (RO)
[29], offering the possibility to inspect the learned model and pro-
viding an estimate of the relative weight of routinely collected
demographic, clinical and biochemical data in predictions. This
model – originally developed for cancer-associated thrombosis risk
assessment [23–25] and then adapted to estimate the risk of dis-
ease progression of breast cancer patients [26] – seeks not only
decision, but also interpretability of the model itself, which repre-
sents the novel aspect of our research. Based on previous observa-
tions, we hypothesized that our previous approach could be of
assistance in predicting the risk of MO in migraine patients. To
achieve this objective, a proof-of-concept study was specifically
designed produce MO risk predictors (RO-MO) using real world
data from a large, well clinical characterized migraine outpatient
population, detailing lifestyle, behavioral and socio-demographic
factors and clinical phenotype. The performance of RO-MO was
also compared with that of baseline SVM to assess the significance
of optimizing the relative importance of groups of clinical attrib-
utes in the selection of MO risk predictors.

2. Patients and methods

2.1. Patient’s dataset

Starting from January 2008, the Headache and Pain Unit of the
Department of Neurological, Motor and Sensorial Sciences and the
InterInstitutional Multidisciplinary Biobank (BioBIM) of the
IRCCS San Raffaele Pisana, Rome, Italy, are jointly involved in the
recruitment of outpatients affected by headache, who are prospec-
tively followed under the appropriate Institutional ethics approval
and in accordance with the principles embodied in the Declaration
of Helsinki [30]. All patients undergo a careful physical and neuro-
ological examination performed by trained neurologists (PB, GE, LF)
and are screened with face-to-face interviews using a semi-
structured questionnaire formulated to collect thorough data on
lifestyle, behavioral and socio-demographic factors (age, sex, civil
status, occupation, body mass index, arterial blood pressure, sport
activity, use of coffee, alcohol, smoking, sleep disturbances, meno-
pause, contraceptive use), comorbidities and concomitant medica-
tions and clinical features of migraine (disease duration, family
history of migraine, presence, frequency, duration of attacks, loca-
tion, accompanying symptoms, unilateral cranial autonomic symp-
toms, triggers and alleviating factors, prodromes, postdromes,
allodynia, presence and duration of medication overuse, current
acute or preventive medications, patient’s satisfaction with tri-
tans) [31]. All patients provide written informed consent, previ-
ously approved by the local Institutional Review Board, to donate
a blood sample to be used in the analysis of possible determinants
of migraine outcome or response to treatment.

Blood samples are withdrawn in fasting conditions and pro-
cessed using standard operating procedures, ICT tools and dedi-
cated software to track the entire sample life, including elapsed
time between blood withdrawal and storage [30]. Routine bio-
chemical analyses are performed on fresh blood samples at time
of enrolment. Thereafter, blood samples are processed, aliquoted,
coded, and stored at –80 °C for subsequent batch analyses. Storage
conditions are carefully monitored, and all aliquots are limited to
one freeze–thaw cycle to ensure the best sample quality. For the present study, 777 consecutive migraine patients were analyzed.

2.2. Experimental settings and statistical analysis

Our MO risk predictor is based on multiple kernel learning over support vector machines and Random Optimization (RO-MO). We built RO-MO in KELP [32] as previously reported [23]. RO-MO was used to produce prognostic discriminators (referred as RO-MO-x) yielding the best classification performance over a training (3-fold cross validation) and test set. Twenty different learning sessions with 20 different RO initializations were performed. The dataset, consisting of 777 patients, was randomly divided into a training and a test set. The training set consisted of 543 migraine patients (70% of the dataset); the remaining 234 patients were allocated to the test set (30% of the cases).

Baseline SVM was run on weka platform and its performance was evaluated in comparison with RO-MO classifiers.

Demographic, clinical and biochemical characteristics of the training and test sets are summarized in Tables 1 and 2.

**Table 1**

| Clinical features of migraine patients in the training and test set. |
|-----------------|-----------------|-----------------|
|                  | Training         | Test            |
| Sex, N (%)       | 0.513            |
| Male             | 96 (17.7)        | 46 (19.7)       |
| Female           | 447 (82.3)       | 188 (80.3)      |
| Menopausal status|                 | 0.235           |
| Pre              | 327 (73.2)       | 146 (77.7)      |
| Post             | 120 (26.8)       | 42 (22.3)       |
| Age at menarche  | 12.2 ± 1.2       | 12.2 ± 1.4      |
| Age (Years), Mean ± SD (range) | 41.5 ± 13.2 (13–78) | 39.5 ± 12.5 (17–71) |
| BMI, Mean ± SD (range) | 23.9 ± 17.0 (170–468) | 23.7 ± 15.0 (150–365) |
| Age of onset (Years), Mean ± SD (range) | 19.9 ± 11.0 (3–73) | 19.9 ± 10.7 (4–62) |
| Length of chronicization (Years), Median (IQR) | 2 (1–4) | 2 (1–5) |
| Type of migraine, N (%) | 0.092           |
| Chronic          | 139 (25.6)       | 48 (20.5)       |
| Episodic without aura | 287 (52.9)        | 114 (48.8)      |
| Menstrual migraine, N (%) | 181 (40.5)          | 92 (39.3)       |
| Episodic with aura | 85 (16.0)        | 54 (23.1)       |
| Cluster headache | 11 (2.0)         | 9 (3.8)         |
| Tension-type headache | 21 (3.9)       | 9 (3.8)         |
| Family, N (%)     |                 | 0.648           |
| Familiarity       | 383 (70.8)       | 168 (72.4)      |
| Frequency (days/months), Median (IQR) | 7 (3–15) | 5 (2–10) |
| Pain Localization, N (%) | 0.377           |
| Unilateral        | 357 (65.8)       | 146 (62.4)      |
| Unilateral or bilateral | 4 (0.7)          | 2 (0.8)         |
| Bilateral         | 182 (33.5)       | 86 (36.8)       |
| Unilateral cranial autonomic symptoms, N (%) | 0.887           |
| Neuropsychiatric  | 167 (30.8)       | 65 (27.8)       |
| Cardiovascular    | 61 (11.3)        | 22 (9.4)        |
| Endocrine-metabolic | 83 (15.4)       | 30 (12.8)       |
| Treatment, N (%)  |                 | 0.009           |
| Type of medication |                 |                 |
| Triptans          | 140 (25.8)       | 71 (30.3)       |
| NSAID             | 165 (30.4)       | 74 (31.6)       |
| Triptans + NSAID  | 178 (32.8)       | 51 (21.9)       |
| Other             | 60 (11.1)        | 38 (16.2)       |
| Response to triptans | 280 (80.0)    | 104 (85.3)      |
| Use of prophylaxis | 304 (56.0)     | 112 (47.9)      |
| Medication overuse|                 | 0.087           |
| Overusing patients, N (%) | 127 (23.4)   | 35 (15.0)       |
| Abused Drug, N (%) |                 | 0.010           |
| Triptans          | 38 (29.9)        | 9 (25.7)        |
| NSAIDs ± combination medications | 19 (15.0) | 7 (20.0) |
| Combination medications | 64 (50.4)   | 15 (42.9)       |
| Others            | 8 (6.3)          | 2 (5.7)         |
| Quantity of abused drug (tablets/month) | 34 ± 29 | 33 ± 26 |
| Overuse duration (Years), Median (IQR) | 2 (1–4) | 2 (1–4) |

BMI: body mass index; NSAID: nonsteroidal anti-inflammatory drugs; IQR: interquartile range.

*p = 0.847 for subgroup analysis.
Receiver operating characteristic (ROC) curve, Bayesian analysis and logistic regression were performed by MedCalc Statistical Software version 13.1.2 (MedCalc Software bvba, Ostend, Belgium) to estimate the probability of medication overuse. All tests were two-tailed and only \( p \)-values lower than 0.05 were regarded as statistically significant.

### 3. Results

Overall, 21% (162 of 777) of the enrolled patients reported the presence of MO which had lasted for at least 2 years (Table 1). No substantial differences were observed for clinical and biomolecular variables between patients included in the training and test sets, with the exception of attack frequency, the use of triptans in combination with NSAID and the percentage of patients with MO, which were all slightly higher in the training compared with the test set.

A set of predictors (named RO-MO-x) was identified using a 3-fold cross validation technique on a training set (n = 543). A test set (n = 234) was used to compute the final performance of risk predictors. As shown in Table 3, RO-MO was capable of improving MO risk prediction compared to baseline SVM as demonstrated by a substantial improvement of the f-measure. Thus, to better characterize the performance of the proposed method, the area under the ROC curve (AUC) and positive and negative likelihood ratios (LRs) were also calculated for all RO-MO models in comparison with baseline SVM. As shown in Table 3, RO-MO-08 was the best performing predictor in the training set with an AUC of 0.79, which was significantly higher than that observed for baseline SVM (\( p = 0.003 \)) (Table 3, Fig. 2) and further increased to 0.81 in the test set (Table 4).

Overall, five RO-MO predictors showed an acceptable predictive performance with AUCs > 0.70 and significant positive and negative LRs (Table 3), but only RO-MO-02, RO-MO-03, RO-MO-08 and RO-MO-12 confirmed their predictive value in the test set (Table 4, Fig. 3). Of interest, all four predictors were not only clinically plausible – as demonstrated by the finding that group 2 variables (migraine clinical features) retained the strongest weight in all models (Table 5) – but they had also a complementary configuration of weights (Table 5, Fig. 4). In particular, RO-MO-02 was weighted on demographics and co-morbidities, RO-MO-03 on DBH polymorphism, RO-MO-08 on treatment details and lifestyle-related trig-

### Table 2

| Blood counts    | Training | Test  | \( p \)-value |
|-----------------|----------|-------|--------------|
| Red blood cells | 4.5 ± 0.5| 4.4 ± 0.42| 0.161        |
| Haematocrit     | 39.4 ± 3.5| 39.4 ± 3.4| 0.811        |
| Hemoglobin      | 13.3 ± 1.3| 13.4 ± 1.2| 0.502        |
| White blood cells| 6.6 ± 1.7| 6.5 ± 1.6| 0.581        |
| Neutrophils     | 3.9 ± 1.4| 3.8 ± 1.2| 0.487        |
| Lymphocytes     | 2.1 ± 0.6| 2.6 ± 0.6| 0.947        |
| Monocytes       | 0.5 ± 0.1| 0.5 ± 0.2| 0.409        |
| Platelets       | 230.2 ± 52.1| 231.7 ± 51.4| 0.818        |
| Mean Platelet Volume | 9.0 ± 1.0| 8.9 ± 1.1| 0.218        |

### Coagulative asset

- International Normalized Ratio: 1.0 ± 0.1, 1.0 ± 0.1, 0.984
- Prothrombin Time: 103.3 ± 13.2, 104.1 ± 16.3, 0.656
- aPTT: 27.9 ± 2.9, 30.3 ± 30.3, 0.228
- Fibrinogen: 282.1 ± 54.5, 279.7 ± 47.6, 0.737

### Glucose metabolic asset

- Fasting blood glucose (mg/dl), Mean ± SD: 82.5 ± 11.5, 84.2 ± 10.0, 0.201
- Fasting insulin (\( \mu \)IU/ml), Mean ± SD: 9.6 ± 9.2, 10.0 ± 9.8, 0.677
- HbA1c (%), Mean ± SD: 5.5 ± 0.3, 5.5 ± 0.4, 0.595

### Blood lipids

- Total cholesterol (mg/dl), Mean ± SD: 202.3 ± 35.7, 195.8 ± 35.3, 0.134
- HDL-cholesterol (mg/dl), Mean ± SD: 56.7 ± 14.0, 56.9 ± 12.9, 0.864
- LDL-cholesterol (mg/dl), Mean ± SD: 125.2 ± 31.4, 120.3 ± 30.7, 0.192
- Triglycerides (mg/dl), Mean ± SD: 102.2 ± 57.2, 93.3 ± 56.5, 0.196

### DBH 19-bp I/D polymorphism

| Allele | Training | Test  | \( p \)-value |
|--------|----------|-------|--------------|
| II     | 130 (46.3) | 45 (42.1) | 0.141        |
| ID     | 60 (21.4)  | 33 (30.8)  | 0.139        |
| DD     | 91 (32.4)  | 29 (27.1)  | 0.139        |

*Available in 388 patients (training set, n = 281; test set, n = 107).
gers and RO-MO-12 on the biochemical and metabolic asset. These findings are consistent with literature data showing that DBH polymorphism is associated with MO in chronic migraine patients [34] and suggest that patients’ metabolic asset, obesity, or lifestyle-related factors might contribute to drug overuse in this clinical setting.

As the performance of RO-MO predictors could be further enhanced, we sought to investigate whether the combination of the best binary predictors (RO-MO-02, RO-MO-03, RO-MO-08 and RO-MO-12) into a decision support system (ML-based DSS) could be of advantage over the individual predictors, or baseline SVM. Using this combined model, we were able to classify migraine patients into five categories: definitely not likely to overuse (0, all predictors negative); probably not likely to overuse (1, at least one positive); possibly of drug overuse (2, three of four positive); definitely at risk of drug overuse (4, all predictors positive).

The discriminatory power of the ML-based DSS was first analyzed in the training set. The results obtained showed an overall improvement of MO risk prediction performance, with an AUC of 0.710 (95% CI: 0.669–0.748), which was higher than that observed with baseline SVM (AUC = 0.710; difference between areas: 0.095, p < 0.001), or each single predictor (RO-MO-2 AUC = 0.726, p < 0.001; RO-MO-3 AUC = 0.717, p < 0.001; RO-MO-8 AUC = 0.787, p = 0.05; RO-MO-12 AUC = 0.737, p < 0.001).

It should be noted that the adoption of a DSS incorporating multiple predictors implies that risk evaluation would be represented by a n-level stratification (generated in the event that risk estimate is achieved by all predictors, or by different combinations). In our model, the use of 4 predictors had led to a 5-level stratification, which was barely comparable with the binary classification of the original models or baseline SVM. Thus, instances were re-coded as 1/0 depending on whether they were positive or not to at least three of the RO-MO predictors (DSS predicted class 3 or 4) and ROC curves were re-analyzed. As reported in Fig. 5, the combined model resulted in an overall improvement of MO risk prediction performance, with a 0.765 AUC, which was higher than that observed with SVM, or each single predictor (Fig. 3).

The improved predictive performance of the ML-based DSS was also confirmed in the test set. As shown in Fig. 6, in fact, the combined model resulted in an overall improvement of MO risk prediction performance, with an AUC of 0.83, a sensitivity and specificity

| Model       | Precision | Recall | F-Measure | AUC (SE) | 95% CI       | +LR     | -LR     |
|-------------|-----------|--------|-----------|----------|--------------|---------|---------|
| SVM Baseline| 0.862     | 0.441  | 0.583     | 0.710 (0.0224) | 0.669–0.748 | 20.4    | 0.57    |
| RO-MO-05    | 0.819     | 0.683  | 0.745     | 0.657 (0.0242) | 0.615–0.697 | 1.95    | 0.53    |
| RO-MO-12    | 0.704     | 0.754  | 0.728     | 0.737 (0.0223) | 0.697–0.773 | 2.77    | 0.35    |
| RO-MO-14    | 0.685     | 0.690  | 0.688     | 0.669 (0.0236) | 0.628–0.708 | 1.93    | 0.47    |
| RO-MO-13    | 0.691     | 0.675  | 0.683     | 0.610 (0.0248) | 0.567–0.651 | 1.57    | 0.64    |
| RO-MO-04    | 0.631     | 0.698  | 0.652     | 0.683 (0.0241) | 0.642–0.722 | 2.91    | 0.55    |
| RO-MO-19    | 0.753     | 0.532  | 0.623     | 0.599 (0.0246) | 0.557–0.641 | 1.80    | 0.73    |
| RO-MO-03    | 0.483     | 0.770  | 0.593     | 0.717 (0.0232) | 0.677–0.754 | 2.67    | 0.41    |
| RO-MO-08    | 0.474     | 0.786  | 0.591     | 0.787 (0.0219) | 0.750–0.821 | 5.26    | 0.34    |
| RO-MO-06    | 0.466     | 0.698  | 0.559     | 0.645 (0.0245) | 0.603–0.685 | 1.89    | 0.57    |
| RO-MO-09    | 0.420     | 0.794  | 0.549     | 0.656 (0.0236) | 0.614–0.696 | 1.80    | 0.49    |
| RO-MO-16    | 0.395     | 0.802  | 0.529     | 0.696 (0.0230) | 0.656–0.735 | 2.18    | 0.41    |
| RO-MO-11    | 0.435     | 0.643  | 0.519     | 0.663 (0.0242) | 0.622–0.703 | 2.77    | 0.60    |
| RO-MO-07    | 0.396     | 0.754  | 0.519     | 0.644 (0.0237) | 0.602–0.684 | 1.69    | 0.51    |
| RO-MO-17    | 0.381     | 0.746  | 0.504     | 0.639 (0.0233) | 0.598–0.680 | 3.28    | 0.68    |
| RO-MO-18    | 0.405     | 0.643  | 0.497     | 0.681 (0.0223) | 0.640–0.721 | 1.89    | 0.39    |
| RO-MO-15    | 0.465     | 0.532  | 0.496     | 0.666 (0.0230) | 0.624–0.705 | 1.81    | 0.44    |
| RO-MO-01    | 0.363     | 0.746  | 0.488     | 0.714 (0.0235) | 0.674–0.752 | 4.71    | 0.52    |
| RO-MO-02    | 0.339     | 0.651  | 0.446     | 0.726 (0.0235) | 0.687–0.763 | 3.94    | 0.47    |
| RO-MO-00    | 0.270     | 0.738  | 0.396     | 0.511 (0.0235) | 0.468–0.554 | 1.07    | 0.97    |
| RO-MO-10    | 0.276     | 0.643  | 0.387     | 0.589 (0.0250) | 0.526–0.631 | 1.57    | 0.74    |
| Average RO-MO | 0.509 | 0.699  | 0.560     | 0.664     | 0.705–0.690 | 3.32    | 0.54    |

AUC: Area Under the Curve; CI: Confidence Interval; LR: Likelihood Ratio; SE: Standard Error.

![Fig. 2. Comparison between the Receiver Operator Characteristics Curves for RO-MO-08 and SVM baseline predictor in the training set.](image-url)
of 0.69 and 0.87, respectively, and an accuracy of 0.87 when MO was predicted by at least three RO-MO models. Logistic regression analysis further confirmed that the ML-based DSS could effectively distinguish between MO and non-MO with ORs of 5.72 (95%CI: 1.56–20.9) and 21.0 (95%CI: 8.5–52) for patients classified as probably (3 RO-MO predictors positive), or definitely at risk of drug overuse (4 RO-MO predictors positive), respectively.

4. Discussion

The present study was designed to investigate the performance of a novel ML-based methodological approach to derive an AI-based DSS for MO risk assessment in migraine patients.

MO is a challenging problem in migraine, mostly chronic migraine, in which is associated with higher pain intensity scores, symptoms severity and disability [5,35], as well as the development of MOH [2,7,8]. MOH is defined as a headache for ≥15 days/month in a patient with a pre-existing primary headache (usually migraine or tension-type headache), and develops as a consequence of regular overuse (more than three months) of simple analgesics such as non-steroidal anti-inflammatory drugs, acetylsalicylic acid and paracetamol for ≥15 days/month, or analgesic combinations, triptans, ergotamine or opioid for ≥10 days/month [2].

Given the high disability rate of MOH and the treatment challenges it poses [7], prevention is especially important in patients

---

Table 4
Analytical performance of machine learning with random optimization in the test set.

| Model        | Precision | Recall | F-Measure | AUC (SE) | 95% CI        | +LR   | -LR  |
|--------------|-----------|--------|-----------|----------|---------------|-------|------|
| SVM Baseline | 0.957     | 0.629  | 0.759     | 0.812 (0.0414) | 0.760–0.863 | 15.0  | 0.34 |
| RO-MO-00     | 0.161     | 0.429  | 0.234     | 0.501 (0.0431) | 0.436–0.567 | 1.01  | 1.00 |
| RO-MO-01     | 0.302     | 0.457  | 0.364     | 0.625 (0.0439) | 0.559–0.687 | 2.65  | 0.71 |
| RO-MO-02     | 0.267     | 0.657  | 0.380     | 0.739 (0.0433) | 0.678–0.794 | 4.17  | 0.44 |
| RO-MO-03     | 0.362     | 0.743  | 0.505     | 0.746 (0.0405) | 0.685–0.800 | 2.96  | 0.34 |
| RO-MO-04     | 0.348     | 0.657  | 0.455     | 0.688 (0.0447) | 0.624–0.747 | 2.92  | 0.53 |
| RO-MO-05     | 0.806     | 0.714  | 0.758     | 0.675 (0.0439) | 0.611–0.735 | 2.14  | 0.49 |
| RO-MO-06     | 0.284     | 0.714  | 0.407     | 0.702 (0.0429) | 0.639–0.760 | 2.44  | 0.44 |
| RO-MO-07     | 0.231     | 0.686  | 0.345     | 0.683 (0.0412) | 0.619–0.742 | 1.97  | 0.41 |
| RO-MO-08     | 0.325     | 0.743  | 0.452     | 0.806 (0.0393) | 0.750–0.855 | 5.69  | 0.30 |
| RO-MO-09     | 0.245     | 0.714  | 0.365     | 0.649 (0.0434) | 0.585–0.710 | 1.77  | 0.51 |
| RO-MO-10     | 0.200     | 0.486  | 0.283     | 0.606 (0.0459) | 0.541–0.669 | 1.71  | 0.70 |
| RO-MO-11     | 0.429     | 0.429  | 0.429     | 0.571 (0.0431) | 0.505–0.635 | 1.71  | 0.82 |
| RO-MO-12     | 0.676     | 0.714  | 0.694     | 0.770 (0.0390) | 0.711–0.822 | 3.34  | 0.30 |
| RO-MO-13     | 0.446     | 0.714  | 0.549     | 0.597 (0.0444) | 0.532–0.661 | 1.42  | 0.64 |
| RO-MO-14     | 0.410     | 0.714  | 0.521     | 0.703 (0.0411) | 0.640–0.761 | 2.21  | 0.39 |
| RO-MO-15     | 0.232     | 0.629  | 0.338     | 0.685 (0.0400) | 0.621–0.744 | 1.92  | 0.38 |
| RO-MO-16     | 0.220     | 0.686  | 0.333     | 0.687 (0.0431) | 0.623–0.746 | 2.20  | 0.46 |
| RO-MO-17     | 0.400     | 0.457  | 0.427     | 0.603 (0.0435) | 0.537–0.666 | 2.24  | 0.75 |
| RO-MO-18     | 0.239     | 0.629  | 0.346     | 0.701 (0.0368) | 0.638–0.759 | 1.94  | 0.30 |
| RO-MO-19     | 0.479     | 0.657  | 0.554     | 0.612 (0.0455) | 0.567–0.694 | 1.86  | 0.62 |
| Average RO-MO | 0.382     | 0.631  | 0.452     | 0.812 (0.045)  | 0.728      | 3.01  | 0.52 |

AUC: Area Under the Curve; CI: Confidence Interval; LR: Likelihood Ratio; SE: Standard Error.

Fig. 3. Comparison of the predictive performance among the analyzed models.
prone to frequent headache and the possibility to predict MO is a compelling challenge in the clinical management of migraine.

Based on previous results [20,23–26], we hypothesized that a combined approach of kernel machines and RO of performance would have found a combination of attributes yielding the best classification performance of MO predictors over a test set. The results obtained demonstrated, for the first time to our knowledge, that this approach can be useful in performing an automated prediction for MO, holding the potential for improving model precision through weighting the relative importance of attributes. The clinical soundness of our approach was also supported by the finding that the best scoring models in terms of both f-measure and AUCs were also clinically plausible, as they were all strongly weighted on clinical features (Group 2), which represent some of the major determinants of MO development identified in epidemiological studies [5]. Furthermore, RO-MO-08, the best performing predictor, was substantially weighted on treatment information (use of prophylaxis, different drug classes, or response to triptans), whose role in the development of MO and MOH is also generally accepted [8]. On the other hand, the DBH 19-bp I/D polymorphism, showed irrelevant weights in three of the four models selected for DSS construction, with the exception of RO-MO-03, which may appear in partial disagreement with our previous finding of a significant association between DBH 19-bp I/D polymorphism and MO in inheritance models of genotype-variable association [34]. However, we must acknowledge that while the latter are designed for statistical inference, ML models are designed for prediction and not to prove relationships within the data.

As stated in Section 1, MO is strongly associated to MOH, although it does not necessarily represent its leading cause [2]. Of interest, some of the RO-MO predictors identified in our analysis also showed a discriminatory power for MOH. MOH was present in 9 of the 234 migraine patients included in the test set, 7 of whom (78%) had a positive RO-MO-12 predicted class. Hence, in this patient subset RO-MO-12 was capable of predicting MOH with a sensitivity and specificity of 89% and 71%, respectively, an AUC of 0.80 and a positive likelihood ratio of 3.1 (95% CI: 2.4–3.9) [data not shown]. Nonetheless, whether this approach might be useful to identify individuals who will likely develop MOH still needs to be explored.

On the other hand, the analysis of clinical/biochemical variables identified other features possibly associated to MO, which were seldom considered before [36–38]. This is the case of biochemical features and lifestyle-related attributes, whose relative weight was considerably represented in RO-MO-08 and RO-MO-12, respectively. Previous studies, in fact, suggested that psychiatric comorbidities (e.g. anxiety disorder or depression) [36], smoking [38], lack of exercise [38] or metabolic syndrome [37] could represent risk factors for the development of MOH. Here, we provide further evidence suggesting that an AI-based approach may exploit significant patterns in data – connoting causality between individual features and MO – that can be used to further explore the pathophysiology of MO and MOH in patients with migraine.

The availability of a set of predictive discriminators further allowed the design of a DSS based on a complementary set of four predictors, to investigate whether a combined approach may be of advantage over individual models. The adoption of a model incorporating four predictors allowed to perform a risk evaluation on a five-level stratification and resulted in an overall improvement of MO risk prediction performance over the single predictors, with an AUC of 0.83 and ORs of 5.7 and 21.0 for patients classified as probably (3 RO-MO predictors positive), or definitely at risk of drug overuse (4 RO-MO predictors positive), respectively. This result could be of considerable importance in migraine, where patient education and counseling to avoid MO remains the single most important way to prevent MOH. The possibility to predict those patients at risk for MO, indeed, would help the neurologist to identify individuals who will likely develop MOH still needs to be explored.
advise people from an early stage and tailor treatment based on individual risk.

The results obtained in the present study are barely comparable with previous studies, as the application of ML to implement medical decision making in migraine has generally addressed the issue of disease classification [13–20,39] or treatment outcome prediction [21,22]. All these studies demonstrated that intelligent systems represent a promising approach for migraine classification, holding potential to revolutionize conventional models of diagnostic deliverance in a contest of personalized medicine. Interestingly, using ML approaches based on magnetic resonance imaging resting-state data [19] or somatosensory evoked potentials [39] it was possible to discriminate migraine patients from healthy subjects with an accuracy of 81% and 88%, respectively. On the other hand, the issue of ML-based risk prediction of MO in migraine has never been addressed before. Even the study by Garcia-Chimeno et al. – reporting an accuracy >90% using SVM – analyzed the possibility to correctly classify between different migraine phenotypes (i.e. episodic vs. chronic migraine with MO) [20]. Conversely, in the present study we specifically addressed the issue of MO risk prediction in patients with various migraine phenotypes, demonstrating that the use of a combined approach of kernel machines and RO of performance outperforms SVM with an accuracy of 87%. This percentage is also higher than that reported by Maizels et al., who demonstrated that the use of a computerized headache assessment tool (CHAT) was able to correctly recognize MO in 82.7% of cases [40].

There are, of course, some limitations to acknowledge. First, the sample size was relatively small, ultimately leading to a small number of recorded events. Second, the model here reported was designed and validated on a dataset derived from a biobank/database project in which all patients are interviewed with a face-to-face semi-structured questionnaire detailing in depth demographic and clinical characteristics. Therefore, the population...
enrolled is highly homogenous and well characterized. Nonetheless, the data here reported demonstrate that the use of ML algorithms and RO models might be of advantage in developing local classifiers capable of predicting MO in migraine, thus supporting the neurologist in the critical phase of migraine clinical and therapeutic decision making. Validation in multicenter prospective studies is needed before making any ML approach into the clinical practice available.

5. Conclusions

In conclusion, a combination of machine learning and random optimization – taking into consideration clinical and biochemical features, drug exposure and lifestyle – might represent a valuable approach to MO prediction in migraine. This is particularly appealing in a context of predictive medicine, in which attributes, routinely collected in electronic health records, may be all used to design new tools for clinical and therapeutic decision making. This approach holds the potential for improving model precision through weighting the relative importance of attributes and demonstrates that variables must be considered in MO risk evaluation, thus strengthening the theory advocated by precision medicine that data should be considered in a more general association, rather than individually.

CRediT authorship contribution statement

Patrizia Ferroni: Conceptualization, Methodology, Writing - original draft. Fabio M. Zanzotto: Conceptualization, Methodology, Software. Noemi Scarpato: Data curation, Methodology, Software. Antonella Spila: Data curation, Visualization. Luisa Fofi: Data curation, Visualization. Gabriella Egeo: Data curation, Visualization. Alessandro Rullo: Visualization, Writing - review & editing. Raffaele Palmiotto: Data curation, Writing - review & editing. Piero Barbanti: Conceptualization, Methodology, Writing - original draft. Fiorella Guadagni: Conceptualization, Supervision, Writing - review & editing.

Declaration of Competing Interest

Piero Barbanti has received consultancy fees from Allergan, Bayer, electroCore, Lusofarmacor, Merck, Visufarma, and advisory fees from TEVA, Novatis and Eli-Lilly. Gabriella Egeo received travel grants and honoraria from Eli-Lilly, Novartis, New Penta and Ecuopharma; Luisa Fofi received travel grants and honoraria from Teva, Eli-Lilly and Novartis; Patrizia Ferroni, Fabio M. Zanzotto, Noemi Scarpato, Antonella Spila, Alessandro Rullo, Raffaele Palmiotto and Fiorella Guadagni report no conflict of interest.

Role of the funding source

This work was partially supported by Grants from the European Social Fund, PNR 2015–2020 ARSO1_01163 PerMedNet under the Italian Ministries of Education, University and Research (CUP B66G18000220005), POR Campania FESR 2014–2020 RARE.PLAT. NET (CUP B63D18000380007) and by the Italian Ministry of Health (institutional funding ricerca corrente). The funding sources had no involvement in the study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

References

[1] Burch R. Migraine and tension-type headache: diagnosis and treatment. Med Clin North Am 2019;103:215–33.
[2] Headache Classification Committee of the International Headache Society (HIS). The International classification of headache disorders, 3rd ed., 2018; Cephalalgia 38: 1–211.
[3] Barbanti P, Egeo G. Pharmacological trials in migraine: it’s time to reappraise where the headache is and what the pain is like. Headache 2011;51:435–41.
[4] Ford JH, Schroeder K, Buse DC, Joshi S, Gelwicks S, et al. Predicting initiation of preventive migraine medications: exploratory study in a large U.S. medical claims database. Curr Med Res Opin 2019. https://doi.org/10.1080/03007995.2019.1655723.
[5] Schwedt TJ, Alain A, Reed ML, Fanning KM, Munjal S, et al. Factors associated with acute medication overuse in people with migraine: results from the 2017 migraine in America symptoms and treatment (MAST) study. J Headache Pain 2018;19:38.
[6] May A, Schulte LH. Chronic migraine: risk factors, mechanisms and treatment. Nat Rev Neurol. 2016;12:455–64.
[7] Chen PK, Wang SJ. Medication overuse and medication overuse headache: risk factors, comorbidities, associated burden and nonpharmaceutical and pharmacologic treatment approaches. Curr Pain Headache Rep 2019;23:60.
[8] Wakerley BR. Medication-overuse headache. Pract Neurol 2019;19:399–403.
[9] Zwart JA, Dyb G, Hagen K, Svebak S, Holmen J. Analgesic use: a predictor of chronic pain and medication overuse headache: the Head-HUNT Study. Neurology 2003;61:160–4.
[10] Katsarava Z, Schneeweiss S, Kurth T, et al. Incidence and predictors for chronicity of headache in patients with episodic migraine. Neurology 2004;62:789–90.
[11] Bigal ME, Serrano D, Buse D, Scher A, Stewart WF, Lipton RB. Acute migraine medications and evolution from episodic to chronic migraine: a longitudinal population-based study. Headache 2008;48:1157–68.
[12] Wang SJ, Puh JL, Lu SR, Juang KD. Outcomes and predictors of chronic daily headache in adolescents: a 2-year longitudinal study. Neurology 2007;68:591–6.
[13] Krawczyk B, Simic D, Simic S, Wozniak M. Automatic diagnosis of primary headaches by machine learning methods. Cent Eur J Med 2013;8:157–65.
[14] Celik U, Yurtay N, Koç ER, Tepe N, Güllüoğlu H, et al. Diagnostic accuracy comparison of artificial immune algorithms for primary headaches. Comput Math Methods Med 2015;2015:465192.
[15] Kriegt R, Ajaaf AJ, Aljumely D, Hussain AJ, Ozge A, et al. An intelligent systems approach to primary headache diagnosis. International conference on intelligent computing; 2017. Dordrecht: Springer.
[16] Vandewiele G, De Backere F, Lannoye K, Vanden Berghe M, Janssens O, et al. A decision support system to follow up and diagnose primary headache patients using semantically enriched data. BMC Med Inform Decis Mak 2018;18:98.
[17] Khayamnia M, Yazdchi M, Heidari A, Foroughipour M. Diagnosis of common headaches using hybrid expert-based systems. J Med Signals Sens 2019:9:174–80.
[18] Yang H, Zhang J, Liu Q, Wang Y. Multimodal MRF-based classification of migraine: using deep learning convolutional neural network. Biomed Eng Online 2018;17:138.
[19] Chong CD, Gaw N, Fu Y, Li J, Wu T, et al. Migraine classification using magnetic resonance imaging resting-state functional connectivity data. Cephalalgia 2017;37:828–44.
[20] Garcia-Chimeno Y, Garcia-Zapirain B, Gomez-Beldarrain M, Fernandez-Ruanoa B, Garcia-Monco JC. Automatic migraine classification via feature selection committee and machine learning techniques over imaging and questionnaire data. BMC Med Inform Decis Mak 2017;17:38.
[21] Parejas Bravo F, Del Barrio Garcia AA, Gallego MM, Gago Vega AB, Ruiz M, et al. Prediction of patient’s response to Onabotulinumtoxin A treatment for migraine. Helyon 2019;5:e01043.
[22] Athreye A, Iyer R, Neavin D, Wang L, Weinsilboum R, et al. Augmentation of preventive migraine medications: exploratory study in a large U.S. medical claims database. Curr Med Res Opin 2019. https://doi.org/10.1080/03007995.2019.1655723.
[23] Gompel A, Alpaydin E. Multiple kernel learning algorithms. J Mach Learn Res 2007;8:217–56.
[24] Cantarini M, Blackman M, Mancini G, Frosini B, Sarti D, et al. Random optimization. Automat Rem Contr 1965;26:246–53.
[25] Barbanti P, Aurilia C, Dal'Arm V, Egeo G, Fofi L, et al. The phenotype of migraine with unilateral cranial autonomic symptoms documents increased peripheral and central trigeminal sensitization. A case series of 757 patients. Cephalalgia 2016;36:1334–40.
[26] Filipic S, Castellucci G, Croce D, Basili R, KeLP: A kernel-based learning platform for natural language processing. Proceedings of ACL-IJCNLP 2015 system demonstrations. Beijing, China, July 26–31; 2015. p. 19–24.
[33] Saar-Tsechansky M, Provost F. Handling missing values when applying classification models. J Mach Learn Res 2007;8:1623–57.
[34] Barbanti P, Guadagni F, De Marchis ML, Ialongo C, Egeo G, et al. Dopamine-beta-hydroxylase 19-bp insertion/deletion polymorphism affects medication overuse in patients with chronic migraine. Neurol Sci 2019;40:1717–24.
[35] Lipton RB, Munjal S, Buse DC, Fanning KM, Bennett A, et al. Predicting inadequate response to acute migraine medication: results from the American migraine prevalence and prevention (AMPP) study. Headache 2016;56:1635–48.
[36] Diener HC, Dodick D, Evers S, Holle D, Jensen RH, et al. Pathophysiology, prevention, and treatment of medication overuse headache. Lancet Neurol 2019;18:891–902.

[37] He Z, Dong L, Zhang Y, Kong Q, Tan G, et al. Metabolic syndrome in female migraine patients is associated with medication overuse headache: a clinic-based study in China. Eur J Neurol 2015;22:1228–34.
[38] Hagen K, Linde M, Steiner TJ, Stovner LJ, Zwart JA. Risk factors for medication-overuse headache: an 11-year follow-up study The Nord-Trøndelag Health Studies. Pain 2012;153:56–61.
[39] Bingzhao Z, Coppola G, Shoaran M. Migraine classification using somatosensory evoked potentials. Cephalalgia 2015;39:1143–55.
[40] Maizels M, Wolfe WJ. An expert system for headache diagnosis: the Computerized Headache Assessment tool (CHAT). Headache 2008;48:72–8.