Open Respiratory Archives xxx (xxxx) xxx–xxx

Sociedad Española de Neumología y Cirugía Torácica SEPAR

Open Respiratory Archives

www.elsevier.es/ora

Original article

# Artificial Intelligence Applied to Forced Spirometry in Primary Care

Rosaly Moreno Mendez[a,*], Antonio Marín[b], José Ramon Ferrando[c], Giuliana Rissi[d], Sonia Cepeda Madrigal[d], Gabriela Agostini[e], Pablo Catalan Serra[a]

[a] Pneumology Department, Arnau de Vilanova Hospital, Valencia, Spain
[b] Data Science in Energy, Av. Eduardo Dato, 22, 41018 Seville, Spain
[c] Pneumology Department, La Ribera Hospital, Alzira, Valencia, Spain
[d] Pneumology Department, Vila-Real Hospital, La Plana, Castellón, Spain
[e] Otolaryngology Department, Vida Clinic, Santa Cruz de Tenerife, Spain

ARTICLE INFO

*Article history:*
Received 17 October 2023
Accepted 19 February 2024
Available online xxx

*Keywords:*
Spirometry
Active case-finding program
Gradient boosting
Cross-validation

ABSTRACT

*Introduction:* This study aims to create an artificial intelligence (AI) based machine learning (ML) model capable of predicting a spirometric obstructive pattern using variables with the highest predictive power derived from an active case-finding program for COPD in primary care.
*Material and methods:* A total of 1190 smokers, aged 30–80 years old with no prior history of respiratory disease, underwent spirometry with bronchodilation. The sample was analyzed using AI tools. Based on an exploratory data analysis (EDA), independent variables (according to mutual information analysis) were trained using a gradient boosting algorithm (GBT) and validated through cross-validation.
*Results:* With an area under the curve close to unity, the model predicted a spirometric obstructive pattern using variables with the highest predictive power: FEV1_theoretical_pre values. Sensitivity: 93%. Positive predictive value: 94%. Specificity: 97%. Negative predictive value: 96%. Accuracy: 95%. Precision: 94%.
*Conclusion:* An ML model can predict the presence of an obstructive pattern in spirometry in a primary care smoking population with no prior diagnosis of respiratory disease using the FEV1_theoretical_pre values with an accuracy and precision exceeding 90%. Further studies including clinical data and strategies for integrating AI into clinical workflow are needed.



## Inteligencia artificial aplicada a la espirometría forzada en Atención Primaria

*Palabras clave:*
Espirometría
Programa de búsqueda activa de casos
Gradiente de aumento
Validación cruzada
Matriz de confusión

RESUMEN

*Introducción:* Este estudio tiene como objetivo crear un modelo de aprendizaje automático (ML) basado en inteligencia artificial (IA) capaz de predecir un patrón obstructivo espirométrico utilizando variables con el mayor poder predictivo derivado de un programa activo de búsqueda de casos de enfermedad pulmonar obstructiva crónica (EPOC) en Atención Primaria.
*Materiales y métodos:* Un total de 1.190 fumadores, de entre 30 y 80 años, sin antecedentes de enfermedad respiratoria, fueron sometidos a espirometría con IA artificial. Sobre la base de un análisis de datos exploratorio (EDA), las variables independientes (según el análisis de información mutua) se entrenaron utilizando un algoritmo de gradiente de aumento (GBT) y se validaron mediante validación cruzada.
*Resultados:* Con un área bajo la curva cercana a la unidad, el modelo predijo un patrón obstructivo espirométrico utilizando los valores del FEV1 prebroncodilatador. Sensibilidad: 93%. Valor predictivo positivo: 94%. Especificidad: 97%. Valor predictivo negativo: 96%. Precisión: 95%. Precisión: 94%.

* Corresponding author.
*E-mail address:* morenomendezrosaly@gmail.com (R.M. Mendez).

https://doi.org/10.1016/j.opresp.2024.100313


Please cite this article as: R.M. Mendez, A. Marín, J.R. Ferrando et al., Artificial Intelligence Applied to Forced Spirometry in Primary Care, Open Respiratory Archives, https://doi.org/10.1016/j.opresp.2024.100313

*Conclusión:* Un modelo ML puede predecir la presencia de un patrón obstructivo en la espirometría en una población fumadora de atención primaria sin diagnóstico previo de enfermedad respiratoria utilizando los valores FEV1 prebroncodilatadores con una exactitud y precisión superiores al 90%. Se necesitan más estudios que incluyan datos clínicos y estrategias para integrar la IA en el flujo de trabajo clínico.


## Introduction

The world's population is aging, leading to an increased prevalence of chronic disorders such as chronic obstructive pulmonary disease (COPD).[1] Early diagnosis and the challenge posed by reducing under diagnosed rates (as well as its classification and treatment) are still pending tasks.[2] Promoting primary prevention and/or providing basic spirometry equipment in primary healthcare centers does not seem to have been turning points in improving this situation. The fact that the disease can manifest at an early age[3] and that interpreting spirometry results is not always straightforward[4] could be variables of interest when planning diagnostic strategies among general practitioners (GPs).

In recent years, the application of artificial intelligence (AI) in the field of medicine has grown exponentially, utilizing various types of data and positively impacting the functional diagnostic accuracy of diseases like COPD[5]. Could clinical decision-making and the automation of healthcare processes find anchor points in machine learning (ML) and deep learning (DL)? The answer is yes, but the data used in the models must adhere to data protection laws,[6] be always accessible, meet a clinical need with relevant outcomes, and undergo exploratory data analysis (EDA). Finally, the validated algorithm must be integrated into the clinical workflow and field management of healthcare centers.[7]

Numerous examples exist of ML integration in the diagnosis of various pulmonary disease.[8–11]

Our goal is to create an AI model based on ML capable of predicting the presence or absence of an obstructive pattern using variables with the highest predictive power derived from an active search program for COPD in primary care.

## Material and methods

### *Patients*

In the period between May 2015 and May 2017, patients referred from six primary care centers in the Valencian Community, Spain, aged between 30 and 80 years with a year-package index equal to or greater than 10, with or without symptoms, were included. Prior diagnosis of respiratory diseases, absence of a signed informed consent, and/or receiving active systemic treatment were considered exclusion criteria.

Forty-four GPs participated in patient inclusion. When a potential candidate patient was identified in the primary care consultation and after the signing of their informed consent, the patient was referred to the spirometry consultation located at each of the health centers in the study area.

### *Spirometry assessment*

All patients underwent forced spirometry and post-bronchodilator test (BDT) in accordance with ATS/ERS guidelines,[12] using the same USB Care Fusion® equipment and trained personnel. Only spirometric assessments quality criteria A and B were analyzed.

### *Variables*

The following variables were analyzed: age, gender, number of cigarettes smoked daily, number of years smoking, forced vital capacity (FVC), forced expiratory volume in 1 second (FEV1) in absolute and theoretical values, FEV1/FVC ratio both before and after BDT, as well as the lower limit of normality (LLN) of the ratio after BDT. The obstructive pattern was defined according to GOLD 2023 consensus criteria.[13]

### *Statistical analysis*

Data was stored and analyzed using the Statistical Package for the Social Sciences (SPSS) version 21.0® (SPSS Inc, Chicago, IL, Estates Unites) (IBM Analytics, Arkoma, NY, EE. UU.).

The statistical analysis initially involved a general descriptive study of the results obtained in all included variables. Results were expressed as mean ± standard deviation for continuous variables (or median and range if the distribution was not normal), and as absolute values and percentages for categorical variables.

### *EDA in Python version 3.8.5*

The data consisted of 1232 rows and 16 columns, including 15 numeric variables and two nominal categorical variables. The target variables were defined as the presence or absence and FEV1/FVC ratio less than 70%. Duplicated, missing, extreme, or atypical values within the dataset were removed.

### *Computerized algorithm and validation*

The development of a computer algorithm for interpreting spirometry results using ML was based on Python version 3.8.5 and the use of mutual information statistics.[14–17]

The importance of variables was estimated using gradient tree boosting (GTB) of LightGBM,[18,19] and a new decision tree based on spirometry data combined with age, gender, and smoking habits was developed. The area under the curve (AUC) was used to evaluate the models. To better assess the model's prediction, a cross-validation 5-fold[20] was performed, where the data was divided into five equal parts, and five iterations were conducted, with each fold used as the validation set (20%) and the remaining as the training set (80%). This technique helps avoid overfitting that could occur with small datasets.

Out-of-fold predictions (off-preds)[20] were used to measure the predictive capability of the model on the already validated data.

The ranking of variables based on their predictive power after the training and validation process was also confirmed using the Explain Like I'm 5 (ELI5) library.[21]

The performance of our classification model was evaluated using a $2 \times 2$ confusion matrix.

During the preparation of this work, the authors used exploratory data analysis (EDA) in Python 3.8.5, gradient boosting algorithm (GBT) of LightGBM, cross-validation 5-fold, out-of-fold predictions (off-preds) and Explain Like I'm 5 (ELI5) to design and validate the machine learning (ML) model. After using these tools,

**Table 1**
Descriptive analysis.

| Patient features | |
|---|---|
| Features | Value |
| Number of patients | 1190 |
| Age | 55.86 ± 10.72 |
| Women | 522 (43.86%) |
| Men | 668 (56.14%) |
| Cigarettes per day | 20.81 ± 11.71 |
| Years of smoking | 31.27 ± 12.11 |
| Pre-quotient FEV1/FVC | 72.89 ± 9.57 |
| Spirometry post-quotient | 74 |
| LLN | 66.02 ± 3.31 |
| FVC_absolute_pre | 3431.86 ± 911.5 |
| FVC_theoretical_pre | 83.16 ± 15.37 |
| FVC_absolute_post | 3444.98 ± 903.04 |
| FVC_theoretical_post | 83.5 ± 15.24 |
| FEV1_absolute_pre | 2511.03 ± 785.81 |
| FEV1_theoretical_pre | 81.62 ± 19.22 |
| FEV1_absolute_post | 2521.12 ± 770.34 |
| FEV1_theoretical_post | 81.89 ± 18.97 |

Pre-quotient: pre-bronchodilator ratio of FEV1 and FVC; Spirometry post-quotient: ratio of FEV1 and FVC after bronchial dilation test; LLN: lower limit of normal; FVC_absolute_pre: pre-bronchodilator forced vital capacity absolute; FVC_theoretical_pre: pre-bronchodilator forced vital capacity theoretical; FVC_absolute_post: post-bronchodilator forced vital capacity absolute; FVC_theoretical_post: post-bronchodilator forced vital capacity theoretical; FEV1_absolute_pre: pre-bronchodilator forced expiratory volume in 1 second absolute; FEV1_theoretical_pre: pre-bronchodilator forced expiratory volume in 1 second theoretical; FEV1_absolute_post: post-bronchodilator forced expiratory volume in 1 second absolute; FEV1_theoretical_post: post-bronchodilator forced expiratory volume in 1 second theoretical.

the authors reviewed and edited the content as needed and assume full responsibility for the publications' content.

The study protocol was approved by the ethics committee of the Arnau de Vilanova-Lliria Hospital located in Valencia, Spain.

## Results

The training dataset included 1190 cases after the completion of EDA. Table 1 provides a descriptive summary of the sample.

### *Predictor variables*

Through mutual information,[14–17] patterns of correlation (dependence) between variables were identified (see Fig. 1).

- *Weak correlations*: among spirometry results, tobacco, gender, and age.
- *Intermediate correlations*: among spirometry-derived results.
- *Strong correlations*: between age and LLN (0.638) and between the pre-BDT ratio and post-BDT ratio, both with the target variable (0.75 and 0.98 respectively).

The positive correlation between pre- and post-BDT spirometry results, along with the more widespread use of forced spirometry without bronchodilator testing in primary care centers, influenced the use of pre-BDT results instead of post-BDT results, without significantly affecting the predictive power of the model based on the AUC of the different classifiers analyzed. Additionally, the other variables with weak and intermediate dependencies were used as input data in the chosen algorithm.

### *ML algorithm with multiple variable combinations*

Using a GTB of LightGBM,[18,19] permuting variables with higher predictive power allowed the analysis of 11 different GTB models, providing an overview of their discriminative capabilities through the resulting AUC values. The standout models were model 2 (mod2), 3 (mod3), and 4 (mod4).

- Mod2: age, gender, pre-quotient (pre-bronchodilator ratio of FEV1 and FVC), FVC_theoretical_pre (pre-bronchodilator forced vital capacity theoretical), FEV1_theoretical_pre (pre-bronchodilator forced expiratory volume in 1 second theoretical).
- Mod3: gender, pre-quotient, FVC_theoretical_pre, FEV1_teórico_pre.
- Mod4: pre-quotient, FVC_theoretical_pre, FEV1_theoretical_pre.

It is worth noting that the AUC value exceeded 0.97 in Mod2, Mod3, and Mod4. Spirometry data alone statistically predict the presence or absence of an obstructive pattern.

### *Classifier model validation*

After using cross-validation 5-fold and off-preds,[20] the most relevant off-preds were displayed (Figs. 2 and 3).

Fig. 2 for the pre-quotient, with both five and three variables, confirmed that the model is accurate, but data dispersion is lower in the model using three features. The predictive probability of the pre-quotient exceeds 0.8 for values below 66%.

The curves in Fig. 3 follow a similar pattern to Fig. 2, with less data dispersion in the case of three variables. However, in the range of 58–82% of the theoretical value of FEV1_theoretical_pre, there was a loss of probabilistic power in the model.

Finally, ELI5[21] ranked pre-quotient and FEV1_theoretical_pre as the variables with the highest probabilistic power.

### *Confusion matrix*

The results are shown in Fig. 4.

*Model A*: Sensitivity (S): 93%. Positive predictive value (PPV): 94%. Specificity (E): 97%. Negative predictive value (PNV): 96%. Precision (P): 94%. Accuracy (Ac): 95%.

*Model B*: S: 73%. PPV: 94%. E: 97%. PNV: 88%. P: 94%. Ac: 90%. The false negative rate increased using the 5-variable model.

## Discussion

The focus was placed on the real contribution of the study on conventional clinical practice.

With the aim of contributing to the diagnosis of COPD, this is the first study that combines AI with spirometry data derived from a case-finding study in primary care. Our study shows that a ML model can predict the presence of an obstructive pattern in spirometry in a primary care population with no prior diagnosis of respiratory disease using the FEV1_theoretical_pre values with an accuracy and precision exceeding 90%.

In the conventional practice of health centers in the Valencian Community, there is no access to spirometry consultations. As a result, in most cases, only those patients with a high symptom burden and frequent visits to the GPs are referred to tertiary care centers with the intention of conducting a complete respiratory functional study and accessing specialized pulmonology consultations. The shortage of resources, both human and material, in health centers hinders the diagnosis of mild cases, patients with few symptoms, young patients, and women. All these factors are considered determinants in the underdiagnosis of COPD.[22,23]

On another note, the use of simpler devices than spirometry, such as COPD-6 among others, allows obtaining FEV1 values at the time of the consultation quickly.[24] The emphasis that our study places on FEV1_theoretical_pre values could contribute to

| id | feature | Age | Gende | Cigarettes per day | Year of smoking | Pre-quotient | Spirometry post-quotient | Lower limit of normal | FVC_absolute_pre | FVC_theoretical_pre | FVC_absolute_post | FVC_theoretical_post | FEV1_absolute_pre | FEV1_theoretical_pre | FEV1_absolute_post | FEV1_theoretical_post | Obstructive pattern |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Age | 1.000 | 0.086 | 0.165 | 0.299 | 0.212 | 0.203 | 0.638 | 0.236 | 0.246 | 0.240 | 0.237 | 0.251 | 0.239 | 0.255 | 0.245 | 0.107 |
| 2 | Gender | 0.017 | 1.000 | 0.023 | 0.022 | 0.014 | 0.015 | 0.021 | 0.044 | 0.020 | 0.045 | 0.019 | 0.034 | 0.019 | 0.032 | 0.019 | 0.026 |
| 3 | Cigarettes per day | 0.096 | 0.069 | 1.000 | 0.100 | 0.082 | 0.079 | 0.071 | 0.100 | 0.090 | 0.099 | 0.095 | 0.104 | 0.098 | 0.098 | 0.097 | 0.027 |
| 4 | Year of smoking | 0.255 | 0.096 | 0.147 | 1.000 | 0.167 | 0.176 | 0.238 | 0.189 | 0.193 | 0.186 | 0.189 | 0.197 | 0.189 | 0.191 | 0.189 | 0.096 |
| 5 | Pre-quotient | 0.168 | 0.058 | 0.112 | 0.156 | 1.000 | 0.474 | 0.117 | 0.172 | 0.179 | 0.169 | 0.180 | 0.207 | 0.247 | 0.205 | 0.242 | 0.757 |
| 6 | Spirometry post-quotient | 0.159 | 0.059 | 0.106 | 0.161 | 0.467 | 1.000 | 0.109 | 0.163 | 0.176 | 0.164 | 0.170 | 0.198 | 0.235 | 0.199 | 0.239 | 0.981 |
| 7 | Lower limit of normal | 0.285 | 0.048 | 0.055 | 0.125 | 0.066 | 0.062 | 1.000 | 0.085 | 0.085 | 0.083 | 0.081 | 0.084 | 0.075 | 0.088 | 0.077 | 0.061 |
| 8 | FVC_absolute_pre | 0.212 | 0.203 | 0.154 | 0.199 | 0.195 | 0.187 | 0.170 | 1.000 | 0.264 | 0.620 | 0.254 | 0.410 | 0.236 | 0.408 | 0.236 | 0.084 |
| 9 | FVC_theoretical_pre | 0.212 | 0.089 | 0.134 | 0.195 | 0.195 | 0.194 | 0.164 | 0.254 | 1.000 | 0.242 | 0.544 | 0.256 | 0.351 | 0.248 | 0.345 | 0.148 |
| 10 | FVC_absolute_post | 0.214 | 0.209 | 0.152 | 0.195 | 0.190 | 0.187 | 0.166 | 0.617 | 0.251 | 1.000 | 0.255 | 0.391 | 0.231 | 0.410 | 0.235 | 0.071 |
| 11 | FVC_theoretical_post | 0.203 | 0.085 | 0.140 | 0.190 | 0.194 | 0.186 | 0.156 | 0.243 | 0.540 | 0.245 | 1.000 | 0.250 | 0.336 | 0.248 | 0.354 | 0.136 |
| 12 | FEV1_absolute_pre | 0.231 | 0.162 | 0.164 | 0.212 | 0.241 | 0.233 | 0.174 | 0.421 | 0.273 | 0.403 | 0.269 | 1.000 | 0.288 | 0.645 | 0.285 | 0.248 |
| 13 | FEV1_theoretical_pre | 0.215 | 0.090 | 0.151 | 0.199 | 0.279 | 0.270 | 0.151 | 0.236 | 0.364 | 0.232 | 0.352 | 0.281 | 1.000 | 0.273 | 0.581 | 0.393 |
| 14 | FEV1_absolute_post | 0.233 | 0.149 | 0.154 | 0.205 | 0.236 | 0.232 | 0.179 | 0.415 | 0.263 | 0.419 | 0.264 | 0.640 | 0.278 | 1.000 | 0.283 | 0.240 |
| 15 | FEV1_theoretical_post | 0.220 | 0.090 | 0.150 | 0.199 | 0.274 | 0.274 | 0.155 | 0.236 | 0.359 | 0.236 | 0.371 | 0.278 | 0.582 | 0.279 | 1.000 | 0.396 |
| 16 | Obstructive pattern | 0.019 | 0.024 | 0.008 | 0.020 | 0.169 | 0.222 | 0.024 | 0.017 | 0.030 | 0.014 | 0.028 | 0.048 | 0.078 | 0.047 | 0.078 | 1.000 |

**Fig. 1.** Mutual information of the recorded variables. Degree of dependence between variables. Pre-quotient: pre-bronchodilator ratio of FEV1 and FVC; Spirometry post-quotient: ratio of FEV1 and FVC after bronchial dilation test; LLN: lower limit of normal; FVC_absolute_pre: pre-bronchodilator forced vital capacity absolute; FVC_theoretical_pre: pre-bronchodilator forced vital capacity theoretical; FVC_absolute_post: post-bronchodilator forced vital capacity absolute; FVC_theoretical_post: post-bronchodilator forced vital capacity theoretical; FEV1_absolute_pre: pre-bronchodilator forced expiratory volume in 1 second absolute; FEV1_theoretical_pre: pre-bronchodilator forced expiratory volume in 1 second theoretical; FEV1_absolute_post: post-bronchodilator forced expiratory volume in 1 second absolute; FEV1_theoretical_post: post-bronchodilator forced expiratory volume in 1 second theoretical. Obstructive pattern.

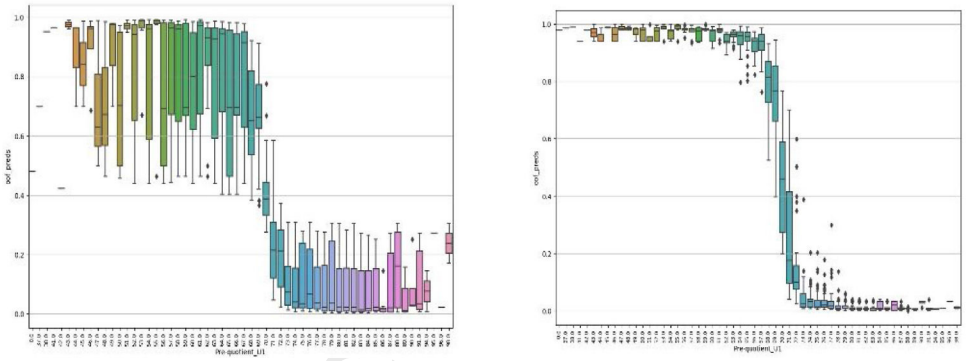

**Fig. 2.** Out-of-fold prediction of the pre-quotient. Left: with five variables (age, gender, pre-quotient, FEV1_theoretical_pre and FVC_theoretical_pre). Right: with three variables (pre-quotient, FEV1_theoretical_pre and FVC_theoretical_pre). Pre-quotient: ratio of FEV1: forced expiratory volume in 1 second. FVC: forced vital capacity. Measured before bronchial dilation test. FEV1_theoretical_pre: pre-bronchodilator forced expiratory volume theoretical in 1 second. FVC_theoretical_pre: pre-bronchodilator forced vital capacity theoretical.

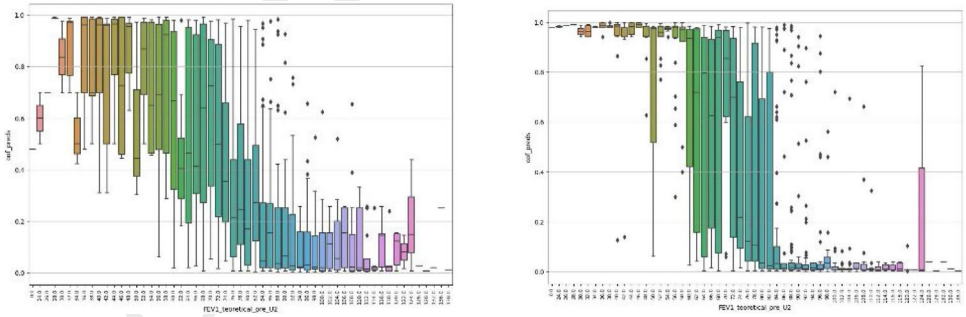

**Fig. 3.** Out-of-fold prediction of the FEV1_pre. Left: with five variables (age, gender, pre-quotient, FEV1_theoretical_pre and FVC_theoretical_pre). Right: with three variables (pre-quotient, FEV1_theoretical_pre and FVC_theoretical_pre). Pre-quotient: ratio of FEV1: forced expiratory volume in 1 second. FVC: forced vital capacity. Measured before bronchial dilation test. FEV1_theoretical_pre: pre-bronchodilator forced expiratory volume theoretical in 1 second. FVC_theoretical_pre: pre-bronchodilator forced vital capacity theoretical.

larger-scale studies among health centers, with the intention of further refining the cutoff points of FEV1_theoretical_pre through a machine learning model capable of reducing overfitting bias. Likewise, it could contribute to future validations of the use of these microspirometers in primary care consultations.

*EDA*

EDA is used as the first step in the data cleaning process.[25] Thanks to this, the ML model uses homogeneous data with the same units and no outliers.[26] In this sense, the data for each patient in

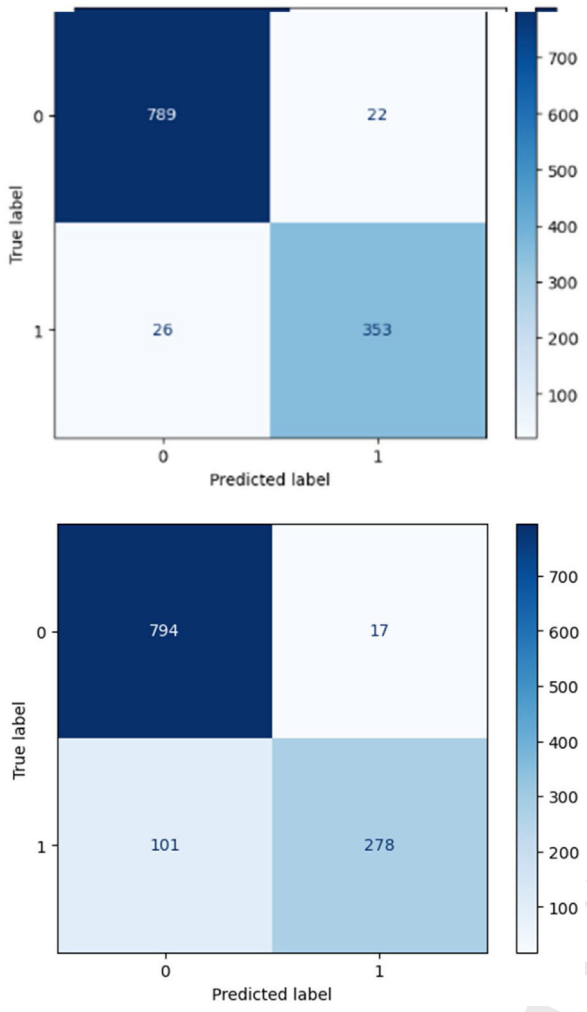


**Fig. 4.** Confusion matrix. Python 3.8.5. A (up): model with three features (pre-quotient, FEV1_theoretical_pre and FVC_theoretical_pre). B (down): model with five features (age, gender, pre-quotient, FEV1_theoretical_pre and FVC_theoretical_pre). Pre-quotient: ratio of FEV1: forced expiratory volume in 1 second. FVC: forced vital capacity. Measured before bronchial dilation test. FEV1_theoretical_pre: pre-bronchodilator forced expiratory volume theoretical in 1 second. FVC_theoretical_pre: pre-bronchodilator forced vital capacity theoretical.

our population was accessible and objective, and it underwent EDA to ensure the quality of our algorithm.

*Classification algorithms: decision trees*

Based on our classification problem, sample size, and the need to handle dimensionality as well as interrelationships between variables, and based on previous literature,[27] our model used decision trees despite the loss of accuracy of this type of algorithm when the disease prevalence is not high.[28] This is a common point with other research groups focused on respiratory pathologies; however, the heterogeneity of the samples in terms of objectives, volume of variables, and analyses used made it difficult to make a comparative analysis between the results of our algorithm and previous groups.[29]

Additionally, the use of modified decision trees as GTB[18,19] allowed individual training of each tree to correct errors made by previous trees, so that they were interconnected and built based on the residual sorted prediction errors of previous trees, gradually reducing the overall error. This allowed us to adapt the model to our positive rate.

*Overfitting of models. Loss of precision*

The advantage of random forests over GTB[18,19] is that the former tolerates overfitting better, meaning the loss of model accuracy when faced with new data. Therefore, in addition to using GTB, our team opted for cross-validation tools previously used.[30,31] Additionally, the use of off-preds[20] provided a more realistic measure of the model's performance on previously unseen data.

After this validation process, our model did not lose statistical or predictive power when using only functional variables (no change in AUC in the absence of information on gender, age, and LLN), or even when using only pre-bronchodilation data. This could be of interest because primary care physicians would use lung values derived from simple devices in their offices to identify patients suspected of having obstructive patterns, where the BDT would eventually be performed with varying speed. Unfortunately, the existing literature that combines case-finding in COPD and the use of AI is scarce. However, we agree with other groups on the relevance of FEV1_theoretical_pre values as a predictor of the presence or absence of obstructive patterns.[28]

*The reality*

For the understanding and using ML models in COPD (a disease with an underlying biological mechanism that is still unknown), the presence of more data (functional, genomic, and clinical) derived from prospective multicenter studies with continuous monitoring is vital.[32]

On the other hand, obtaining optimal metrics does not automatically guarantee a positive impact, so paired studies comparing AI with conventional practice and the integration of predictions into healthcare workflows are required. Ethical questions about the use of AI, such as assigning responsibility in the case of an incorrect diagnosis or misuse of the model, also need to be addressed. Multidisciplinary committees are necessary to ensure effective and safe implementation.[6,22,33,34]

*Limitations*

Our study has several limitations, both stemming from our methodology and from the evaluation of the utility of AI in conventional clinical care.

In the first case, out study used a limited number of patients; however, the AI tools we used allowed us to avoid model overfitting under these circumstances. Another limitation was the lack of social and clinical data; however, the initial goal of our study was to relay on rapidly accessible data in primary care consultations to ensure that the model could predict the presence or absence of an obstructive pattern.

In the second case, as integrating AI into healthcare workflows is an inevitable challenge, we must find integration pathways through tools that are already being used, such as apps among students and doctors. This will allow us to obtain more multicenter, functional, genomic, and social data. With these data, trials, paired studies, and real-time studies can be conducted to increase the reliability of AI as support for our work rather than as an adversary.

## Conclusion

Based on our results, a ML model with GBT is capable of predicting the presence of an obstructive pattern in spirometry, using the pre-bronchodilation FEV1 value as a predictor variable, in a population of primary care smokers without a prior diagnosis of respiratory disease. Further studies that include clinical and longitudinal data are needed, as well as strategies for integrating AI into healthcare workflows. It is our duty to harness the incredible

resources of AI to benefit the millions of people who currently suffer from and will suffer from COPD.[35] This is what we must continue to do despite the limited training in AI tools in medical schools. The time for multidisciplinary teams with data experts and healthcare professionals has arrived.

## Funding

This study was carried out through a grant offered by Boehringer Ingelheim. The funding for the grant was managed by The Foundation for the Advancement of Health and Biomedical Research in the Valencian Community (FISABIO).

## Authors' contributions

Moreno MD is the main researcher and writer of the manuscript.

Ferrando, Rissi, Cepeda, Agostini MD and Catala MD PhD have reviewed the manuscript.

Marin has conducted the statistical analysis and code writing.

Data acquisition and writing the manuscript: RM, AM, JF, GR, SC, GA, PC.

## Conflicts of interest

The authors declare no conflict of interest.

## References

1. Burney P, Jithoo A, Kato B, Janson C, Mannino D, Nizankowska-Mogilnicka E, et al. Chronic obstructive pulmonary disease mortality and prevalence: the associations with smoking and poverty – a BOLD analysis. Thorax. 2014;69:465–73, http://dx.doi.org/10.1136/thoraxjnl-2013-204460.
2. Luis Izquierdo J, Casanova C, Celli B, Santos S, Sibila O, Sobradillo P, et al. The 7 cardinal sins of COPD in Spain. Arch Bronconeumol. 2021;58:498–503, http://dx.doi.org/10.1016/j.arbres.2021.12.008.
3. Represas-Represas C, Botana-Rial M, Leiro-Fernández V, González-Silva AI, García-Martínez A, Fernández-Villar A. Short- and long-term effectiveness of a supervised training program in spirometry use for primary care professionals. Arch Bronconeumol. 2013;49:378–82, http://dx.doi.org/10.1016/j.arbres.2013.01.001.
4. Miravitlles M, Soriano JB, García-Río F, Muñoz L, Duran-Tauleria E, Sanchez G, et al. Prevalence of COPD in Spain: impact of undiagnosed COPD on quality of life and daily life activities. Thorax. 2009;64:863–8, http://dx.doi.org/10.1136/thx.2009.115725.
5. Topalovic M, Das N, Burgel P-R, Daenen M, Derom E, Haenebalcke C, et al. Artificial intelligence outperforms pulmonologists in the interpretation of pulmonary function tests. Eur Respir J. 2019;53, http://dx.doi.org/10.1183/13993003.01660-2018 [Article 1801660].
6. Lillywhite A, Wolbring G. Coverage of ethics within the artificial intelligence and machine learning academic literature: the case of disabled people. Assist Technol. 2021;33:129–35, http://dx.doi.org/10.1080/10400435.2019.1593259.
7. Ben-Israel D, Jacobs WB, Casha S, Lang S, Ryu WHA, de Lotbiniere-Bassett M, et al. The impact of machine learning on patient care: a systematic review. Artif Intell Med. 2020;103:101785, http://dx.doi.org/10.1016/j.artmed.2019.101785.
8. Gonem S, Janssens W, Das N, Topalovic M. Applications of artificial intelligence and machine learning in respiratory medicine. Thorax. 2020;75:695–701, http://dx.doi.org/10.1136/thoraxjnl-2020-214556.
9. Burki TK. Predicting lung cancer prognosis using machine learning. Lancet Oncol. 2016;17:e421, http://dx.doi.org/10.1016/S1470-2045(16)30436-3.
10. Min X, Yu B, Wang F. Predictive modeling of the hospital readmission risk from patients' claims data using machine learning: a case study on COPD. Sci Rep. 2019;9:2362, http://dx.doi.org/10.1038/s41598-019-39071-y.
11. Barton C, Chettipally U, Zhou Y, Jiang Z, Lynn-Palevsky A, Le S, et al. Evaluation of a machine learning algorithm for up to 48-hour advance prediction of sepsis using six vital signs. Comput Biol Med. 2019;109:79–84, http://dx.doi.org/10.1016/j.compbiomed.2019.04.027.
12. Brusasco V, Crapo R, Viegi G, American Thoracic Society European Respiratory Society. Coming together: the ATS/ERS consensus on clinical pulmonary function testing. Eur Respir J. 2005;26:1–2, http://dx.doi.org/10.1183/09031936.05.00034205.
13. Agusti A, Böhm M, Celli B, Criner GJ, Garcia-Alvarez A, Martinez F, et al. GOLD COPD DOCUMENT 2023: a brief update for practicing cardiologists. Clin Res Cardiol. 2023:1–10, http://dx.doi.org/10.1007/s00392-023-02217-0.
14. Baudot P, Tapia M, Bennequin D, Goaillard J-M. Topological information data analysis. Entropy. 2019;21:869, http://dx.doi.org/10.3390/e21090869.
15. Estimating entropy and mutual information with scikit-learn [WWW Document]. Gist. https://gist.github.com/nvictus/c3d336b184fd49ab4553904c74032d90 [accessed 26.08.23].
16. Non-parametric computation of entropy and mutual-information [WWW Document]. Gist. https://gist.github.com/shurain/09421ae79ce81e67060a [accessed 26.08.23].
17. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. Phys Rev E. 2004;69:066138, http://dx.doi.org/10.1103/PhysRevE.69.066138.
18. Welcome to LightGBM's documentation! – LightGBM 3.3.2 documentation [WWW Document], n.d. https://lightgbm.readthedocs.io/en/v3.3.2/ [accessed 17.08.23].
19. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: a highly efficient gradient boosting decision tree. In: Proceedings of the 31st international conference on neural information processing systems NIPS'17. 2017. p. 3149–57.
20. Sohil F, Sohali MU, Shabbir J. An introduction to statistical learning with applications in R: by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, New York Springer Science and Business Media, 2013, $41.98, eISBN: 978-1-4614-7137-7. Stat Theory Relat Fields. 2022;6:87, http://dx.doi.org/10.1080/24754269.2021.1980261.
21. Overview – ELI5 0.11.0 documentation [WWW Document], n.d. https://eli5.readthedocs.io/en/latest/overview.html [accessed 17.08.23].
22. Ancochea J, Miravitlles M, García-Río F, Muñoz L, Sánchez G, Sobradillo V, et al. Infradiagnóstico de la enfermedad pulmonar obstructiva crónica en mujeres: cuantificación del problema, determinantes y propuestas de acción. Arch Bronconeumol. 2013;49:223–9, http://dx.doi.org/10.1016/j.arbres.2012.11.010.
23. Soriano JB, Alfageme I, Miravitlles M, De Lucas P, Soler-Cataluña JJ, García-Río F, et al. Prevalence and determinants of COPD in Spain: EPISCAN II. Arch Bronconeumol. 2021;57:61–9, http://dx.doi.org/10.1016/j.arbres.2020.07.024.
24. Toda R, Hoshino T, Kawayama T, Imaoka H, Sakazaki Y, Tsuda T, et al. Validation of "Lung Age" measured by spirometry and handy electronic FEV1/FEV6 meter in pulmonary diseases. Intern Med. 2009;48:513–21, http://dx.doi.org/10.2169/internalmedicine.48.1781.
25. Cohen JP, Cao T, Viviano JD, Huang C-W, Fralick M, Ghassemi M, et al. Problems in the deployment of machine-learned models in health care. CMAJ. 2021;193:E1391–4, http://dx.doi.org/10.1503/cmaj.202066.
26. Beam AL, Kohane IS. Big data and machine learning in health care. JAMA. 2018;319:1317–8, http://dx.doi.org/10.1001/jama.2017.18391.
27. Bi Q, Goodman KE, Kaminsky J, Lessler J. What is machine learning? A primer for the epidemiologist. Am J Epidemiol. 2019;188:2222–39, http://dx.doi.org/10.1093/aje/kwz189.
28. Spathis D, Vlamos P. Diagnosing asthma and chronic obstructive pulmonary disease with machine learning. Health Informatics J. 2019;25:811–27, http://dx.doi.org/10.1177/1460458217723169.
29. Mekov E, Miravitlles M, Petkov R. Artificial intelligence and machine learning in respiratory medicine. Expert Rev Respir Med. 2020;14:559–64, http://dx.doi.org/10.1080/17476348.2020.1743181.
30. Feng Y, Wang Y, Zeng C, Mao H. Artificial intelligence and machine learning in chronic airway diseases: focus on asthma and chronic obstructive pulmonary disease. Int J Med Sci. 2021;18:2871–89, http://dx.doi.org/10.7150/ijms.58191.
31. Topalovic M, Laval S, Aerts J-M, Troosters T, Decramer M, Janssens W. Automated interpretation of pulmonary function tests in adults with respiratory complaints. Respir Int Rev Thorac Dis. 2017;93:170–8, http://dx.doi.org/10.1159/000454956.
32. Agusti A, Calverley PMA, Celli B, Coxson HO, Edwards LD, Lomas DA, et al. Characterisation of COPD heterogeneity in the ECLIPSE cohort. Respir Res. 2010;11:122, http://dx.doi.org/10.1186/1465-9921-11-122.
33. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. JAMA. 2013;20:144–51, http://dx.doi.org/10.1136/amiajnl-2011-000681.
34. Lovejoy CA, Phillips E, Maruthappu M. Application of artificial intelligence in respiratory medicine: has the time arrived? Respirol Carlton Vic. 2019;24:1136–7, http://dx.doi.org/10.1111/resp.13676.
35. Celli B. From Laennec's stethoscope to the magic of imaging big data and artificial intelligence: a timeline of precision medicine for patients with COPD. Am J Respir Crit Care Med. 2023;208:342–4, http://dx.doi.org/10.1164/rccm.202303-0550ED.