International Journal of Chronic Obstructive Pulmonary Disease Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Curve-Modelling and Machine Learning for a Better COPD Diagnosis

Adriana Maldonado-Franco[1], Luis F Giraldo-Cadavid[2,3], Eduardo Tuta-Quintero[2], Mauricio Cagy[4], Alirio R Bastidas Goyes[2], Daniel A Botero-Rosas[2]

[1]School of Engineering, Universidad de La Sabana, Chía, Colombia; [2]School of Medicine, Universidad de La Sabana, Chía, Colombia; [3]Interventional Pulmonology Service, Fundación Neumológica Colombiana, Bogotá, DC, Colombia; [4]Biomedical Engineering Program, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil

Correspondence: Daniel A Botero-Rosas, Universidad de La Sabana, Morphophysiology Department, Km 7, Northern Highway, Chía, Cundinamarca, 140013, Colombia, Email Daniel.botero@unisabana.edu.co

Background: Development of new tools in artificial intelligence has an outstanding performance in the recognition of multi-dimensional patterns, which is why they have proven to be useful in the diagnosis of Chronic Obstructive Pulmonary Disease (COPD).
Methods: This was an observational analytical single-centre study in patients with spirometry performed in outpatient medical care. The segment that goes from the peak expiratory flow to the forced vital capacity was modelled with quadratic polynomials, the coefficients obtained were used to train and test neural networks in the task of classifying patients with COPD.
Results: A total of 695 patient records were included in the analysis. The COPD group was significantly older than the No COPD group. The pre-bronchodilator (Pre BD) and post-bronchodilator (Post BD) spirometric curves were modelled with a quadratic polynomial, and the coefficients obtained were used to feed three neural networks (Pre BD, Post BD and all coefficients). The best neural network was the one that used the post-bronchodilator coefficients, which has an input layer of 3 neurons and three hidden layers with sigmoid activation function and two neurons in the output layer with softmax activation function. This system had an accuracy of 92.9% accuracy, a sensitivity of 88.2% and a specificity of 94.3% when assessed using expert judgment as the reference test. It also showed better performance than the current gold standard, especially in specificity and negative predictive value.
Conclusion: Artificial Neural Networks fed with coefficients obtained from quadratic and cubic polynomials have interesting potential of emulating the clinical diagnostic process and can become an important aid in primary care to help diagnose COPD in an early stage.
Keywords: artificial neural networks, machine learning, diagnosis, accuracy, COPD

Introduction

Chronic Obstructive Pulmonary Disease (COPD) is a respiratory condition that causes persistent airflow obstruction, due to alterations in the airways and/or in the alveoli, which result in chronic respiratory symptoms such as cough and dyspnoea.[1] Even though COPD is a preventable and treatable disease, it is the third leading cause of death worldwide, with a higher burden on low- and medium-income countries.[2] For example, in Latin America, the PLATINO Study found a 14.3% COPD prevalence, ranging from 7.8% in Mexico City to 19.7% in Montevideo, Uruguay.[3] Globally, more than 3 million deaths were attributed to COPD in 2019 alone,[4] and the disease prevalence is only expected to increase in the following years, because of the world's population aging and continuous exposure to noxious particles or gases, like tobacco smoking and air pollution.[5] Due to its progressive nature, timely diagnosis of COPD is key to manage disease development by taking protective measures and treating symptoms, improving the prognosis for the patient.[6]

Current international guidelines indicate that any subject exposed to at least one risk factor should be tested for COPD with spirometry, and the ratio between the Forced Expiratory Volume in the first second (FEV_1) and the Forced Vital Capacity (FVC), after applying a bronchodilator, is evaluated. If this ratio results to be under 0.7 (70%), the patient is diagnosed with COPD.[1] However, authorities like the American Thoracic Society (ATS) and the European Respiratory

Received: 25 January 2024
Accepted: 20 May 2024
Published: 12 June 2024

Society (ERS) advise considering the statistically derived lower limit of normal (LLN) as well.[7] Even though there has been a global effort in improving its diagnosis, COPD still has high underdiagnosis and misdiagnosis rates. For instance, the PUMA Study, which was conducted in Argentina, Colombia, Uruguay, and Venezuela in 2016, found an underdiagnosis prevalence of 77% and a misdiagnosis rate of 30.4%.[8]

The FEV_1/FVC ratio only considers two instantaneous measures of the spirometry, but more information can be gathered from the flow-volume curve shape. Previous studies have been conducted to quantify and analyse the curvature of this trace with promising results in COPD diagnosis.[9–14] Meanwhile, other researchers have leveraged the increase in computational power and the development of new tools in artificial intelligence. These tools have an outstanding performance in recognizing multidimensional patterns, so they have been proved to be useful in diagnosing respiratory diseases.[15–20] We developed a new COPD diagnostic system that combines these two approaches by modelling the curvature of the expiratory flow-volume trace and feeding the resulting coefficients to different machine-learning techniques to classify patients between those with COPD and those without it. This paper presents the results of the diagnostic accuracy of the proposed system, when tested in patients who are >40 years old and have respiratory symptoms and/or risk factors, in comparison to the classification made by an expert pneumologist.

Methods

Study Design

This was an observational analytical single-centre study that followed the Declaration of Helsinki and was approved by the Ethics Committee of Clínica Universidad de la Sabana. These patients had signed an Informed Consent, in which they authorized the use of their personal data for research purposes. Their records were collected between August 2017 and August 2019.

Eligibility Criteria

The inclusion criteria were patients included in the spirometer database, with signed informed consent and under study due to suspicion of respiratory disease. The exclusion criteria were patients younger than 40 years, having less than 3 post-bronchodilator spirometric trials, incomplete clinical information, or illegible spirometry according to the expert pneumologist. In particular, the record needed to have all the clinical variables that were deemed relevant for COPD diagnosis in a previously conducted expert consensus Supplementary Table 1.

Variables

All included patients had raw spirometry data stored in the same spirometer (Vmax Encore 22, CareFusion, Yorba Linda, California), and demographic information and answers regarding risk factors and respiratory symptoms were collected through questionnaires validated for COPD diagnosis. Spirometry raw data were stored in .fvl format, and each file included the patient's name and ID, date of test, as well as the registration of flow and volume in time for up to 8 pre-bronchodilator and 8 post-bronchodilator forced-manoeuvre spirometry trials. Each file also had the selection of the best pre-bronchodilator trial and the best post-bronchodilator trial, according to a professional in respiratory therapy, who chose the best traces according to the ATS/ERS criteria (the curve with the largest sum of FVC + FEV1 that fulfilled the Acceptability, Usability, and Repeatability ATS/ERS Criteria for FEV1 and FVC).[21] An expert pneumologist evaluated each patient's record and determined which patients had COPD and which did not. This classification was used as a reference test to evaluate diagnostic accuracy.

Study Sample Size

Since this is a study that compares two diagnostic tests, the sample size calculation was performed as indicated by Machin et al.[22] With a significance level of 5%, a power of 20%, an increase in sensitivity from 80% in the imperfect gold-standard to 90% in the new test, and a proportion of diseased patients of 30%, we obtained a minimal sample size of 657.

Test Methods

We extracted the raw signals from the spirometer for all the patients, verified the fulfilment of all inclusion/exclusion criteria, selected the best prebronchodilator and postbronchodilator trial traces as tagged by the respiratory therapist and modelled the expiratory phase from the peak expiratory flow to the forced vital capacity, by means of four techniques:

- quadratic polynomial: $F = C_2V^2+C_1V+C_0$
- cubic polynomial: $F = C_3V^3+C_2V^2+C_1V+C_0$
- one-term exponential: $F = ae^{bV}$
- two-term exponential: $F = ae^{bV}+ce^{dV}$

The goodness of fit for the obtained models was calculated separately for prebronchodilator and postbronchodilator traces by four metrics: Sum Squared Error (SSE), Coefficient of Determination (R^2), adjusted Coefficient of Determination ($adjR^2$) and Root Mean Square Error (RMSE).

Then, the subjects were randomly assigned to training or testing, in a 90/10 ratio.[23] This was done to set aside a group of samples (testing) that will not be used to train the machine learning techniques. This allows the evaluation of the performance of the system in the samples used for training and in samples the system has not seen before, to emulate what would occur in a real-life setting. The coefficients of training samples were fed in three groups (prebronchodilator coefficients, postbronchodilator coefficients and all coefficients) to 5 different machine learning techniques: two unsupervised techniques (k-means clustering and hierarchical clustering) and three supervised techniques (Decision Trees – DT, Support Vector Machines – SVM and artificial neural networks – ANN). The artificial neural networks to be tested were fully connected feedforward networks, and the input layer changed depending on the number of coefficients chosen to feed it. The training of all 5 techniques was performed using 10-fold cross-validation.[24]

Next, the classification made by these techniques on both training and testing groups was compared to the classification made by the expert. The diagnostic performance of each of the machine learning methods was determined based on the accuracy, sensitivity, specificity, positive predictive value, and negative predictive value calculated from the confusion matrix.[25]

Statistical Methods

Categorical Variables were characterized by absolute and relative frequencies. Quantitative variables were described by medians and interquartile ranges if they were not normal according to the One-sample Kolmogorov–Smirnov test. To compare the groups of patients with COPD and without COPD, we used Chi-Squared Pooled Estimate of Proportion for categorical variables and Wilcoxon Rank Sum Test in quantitative variables.

Data Analysis and Modelling Software

Data were processed using MATLAB Release 2022a (The MathWorks, Inc., Natick, Massachusetts, United States).

Results

The available database had spirometric records and questionnaire answers of 765 patients. After reviewing inclusion/exclusion criteria, 695 patient records were included in the analysis Figure 1.

The 695 subjects were divided into two groups (COPD and No COPD), according to the diagnosis made by the expert, and 34% (237/695) had COPD Table 1. The COPD group was significantly older than the No COPD group, and women were the majority in the No COPD group. Most of the subjects (595/695) reported having respiratory symptoms. Out of the whole sample, 27% of the subjects informed having a previous COPD diagnosis, 14% stated they had been diagnosed with asthma, but less than half (46%) claimed they had gone through a spirometry study before. The COPD group had a higher proportion of smokers or ex-smokers (52% vs 42% in the No COPD group), and dyspnoea was the most frequently reported symptom (62%). Also, in the COPD group, the most common risk factor was exposure to wood smoke (68%), followed by tobacco smoking (52%), with a median pack-year of 29, without considering non-smokers.

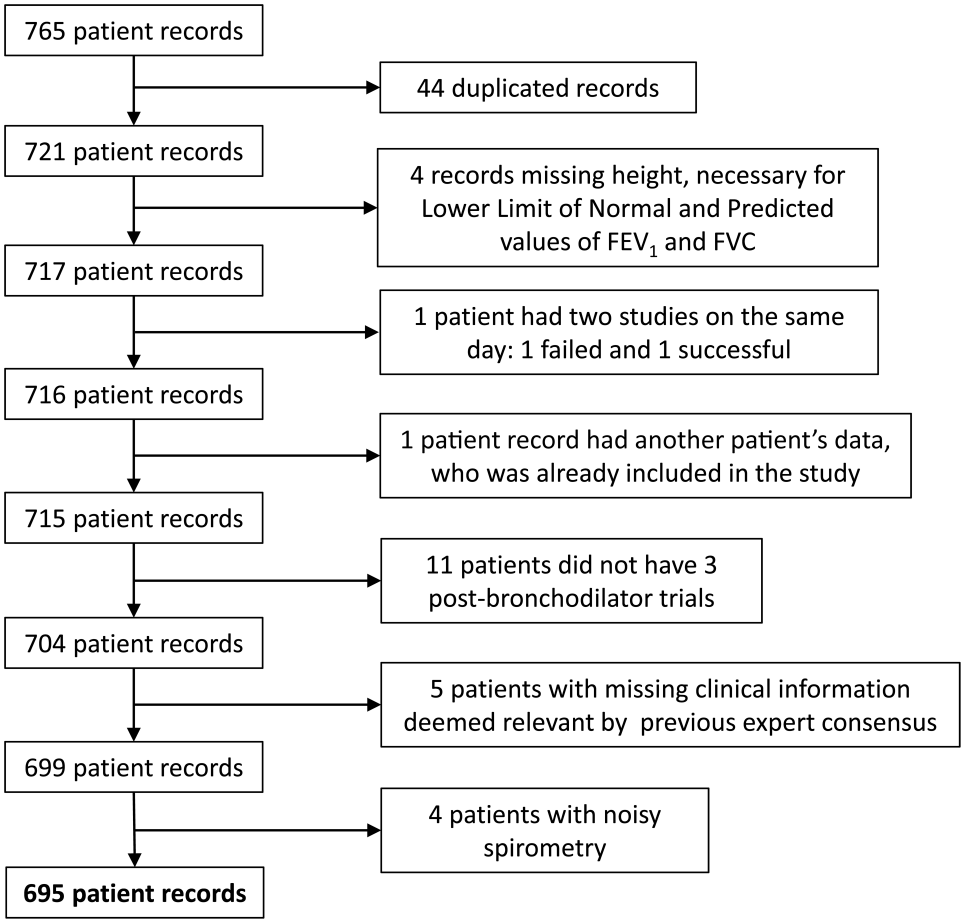

Figure 1 Patient selection process.

Figure 2 shows an example of each of the techniques used for curve fitting. The spirometric curve was modelled from the Peak Expiratory Flow (PEF), or maximum flow, to the Forced Vital Capacity (FVC), or maximum volume.

Table 2 displays the summary of the goodness-of-fit evaluation for each of the modelling techniques, discriminated on prebronchodilator (PreBD) and postbronchodilator (PostBD) traces. The cubic polynomial had the highest goodness-of-fit

Table 1 Patient Demographic Data

	COPD n = 237	No COPD n = 458
Age in years, *m*(IQR)	71 (63–78)	63 (55–71)*
Women, n (%)	112 (47)	280 (61)*
Occupation - Homemaker, n (%)	88 (37)	175 (38)
Smoker/Ex-Smoker, n (%)	124 (52)	192 (42)*
Years Smoking (Without Non-Smokers), *m*(IQR)	29 (13–41)	20 (10–33)*
Daily Cigarettes (Without Non-Smokers), *m*(IQR)	9 (3–20)	4 (2–10)*
Pack-Year (Without Non-Smokers), *m*(IQR)	10 (3–30)	4.98 (1.65–14.2)*
Passive Smoker, n (%)	42 (18)	100 (22)
Daily Cigarettes of Other Smoker (Only Passive Smokers), *m*(IQR)	11 (4–20)	10 (5–20)
Wood Smoke Exposure, n (%)	160 (68)	249 (54)*
Years of Wood Smoke Exposure, *m*(IQR)	20 (10–30)	17 (10–25)
Daily Hours of Wood Smoke Exposure, *m*(IQR)	5 (3–9.5)	6 (3–10)

(*Continued*)

Table 1 (Continued).

	COPD n = 237	No COPD n = 458
Previous Diagnosis of COPD, Chronic Bronchitis, or Emphysema, n (%)	100 (42)	88 (19)*
Previous Diagnosis of Asthma, Asthmatic, or Allergic Bronchitis, n (%)	38 (16)	58 (13)
Previous Spirometry, n (%)	131 (55)	186 (41)*
Presence of Respiratory Symptoms, n (%)	209 (88)	386 (84)
Years With Symptoms, *m*(IQR)	6 (2–20.75)	4 (2–12)*
History of Atopy, n (%)	61 (26)	123 (27)
Presence of Wheezing, n (%)	94 (40)	112 (24)*
Frequency of Wheezing - Sometimes, n (%)	72 (30)	96 (21)*
Presence of Dyspnoea, n (%)	148 (62)	261 (57)
Frequency of Dyspnoea During Physical Activity - Very Frequent, n (%)	65 (27)	89 (19)*
Presence Of Chronic Cough, n (%)	116 (49)	204 (45)
Chronic Cough in The Morning, n (%)	92 (39)	126 (28)*
Chronic Expectoration, n (%)	79 (33)	118 (26)*

Notes: * p<0.05 Wilcoxon Rank Sum Test or Chi Squared Pooled Estimate of Proportion, as appropriate.
Abbreviations: COPD, Chronic Obstructive Pulmonary Disease; m, median; IQR, interquartile range; n, number.

metrics, while the 1-term exponential had the worse fit performance. Since both SSE and RMSE are measures of error, the smaller they are, the better the fit. On the other hand, R^2 and adjR^2 are unity bound and a measure of one means perfect fit.

Table 3 compares, in terms of median and interquartile range, the behaviour of the coefficients between the COPD and No COPD groups, and it shows that most of the proposed coefficients have statistically significant differences between both groups.

The methods that displayed the highest accuracy (<85%) when evaluated in the test samples are summarized in Table 4. The highest accuracy was obtained by the Artificial Neural Network, fed by the coefficients of the quadratic polynomial model applied to the postbronchodilator traces, with an accuracy of 92.9%.

Out of this selection of well-performing methods, this neural network configuration also displayed the highest sensitivity (88.2%) and highest specificity (94.3%). The architecture of this ANN is a 3-neuron input layer and three hidden layers (18, 272 and 1 neurons, respectively), both with hyperbolic tangent as activation function. The output layer has two neurons, one for each class, and the activation function is softmax. Full Results regarding diagnostic accuracy are in the Supplementary Tables 2.1–2.10

For comparison, we analysed the classification made by the current gold standard (FEV1/FVC <0.7) on the 70 patients that were chosen for testing the artificial neural network. Table 5 displays the confusion matrix for the gold standard classification for the testing samples, using the best PostBD curve. Table 6 displays the classification made by the artificial neural network trained with the coefficients from the quadratic polynomial in the same curve and in the same samples.

These two classifications yield the diagnostic accuracy metrics that are summarized in Table 7.

Discussion

The goal of this study was to develop an automated tool that leverages the information contained in the curvature of the spirometry traces to assist the process of diagnosing COPD, particularly in primary care settings, where medical professionals might not be as familiarized with spirometric patterns as specialists. We proposed a new COPD diagnostic system that models the curvature of the expiratory flow-volume trace through a mathematical model and feeds the coefficients of the model to a machine-learning technique to classify patients between those with COPD and those without it. We evaluated the combination of four mathematical models and five machine learning techniques. Out of the different combinations, the best diagnostic performance was obtained with a quadratic polynomial and an artificial neural network.

From the goodness-of-fit scores, we found that cubic polynomial model had the best fit to the spirometric traces, while the 1-term exponential model had the worst. Nonetheless, we chose to use the coefficients from all four models because they all had an acceptable fit to the curves. Then, to get an idea of the proposed variables' capabilities to

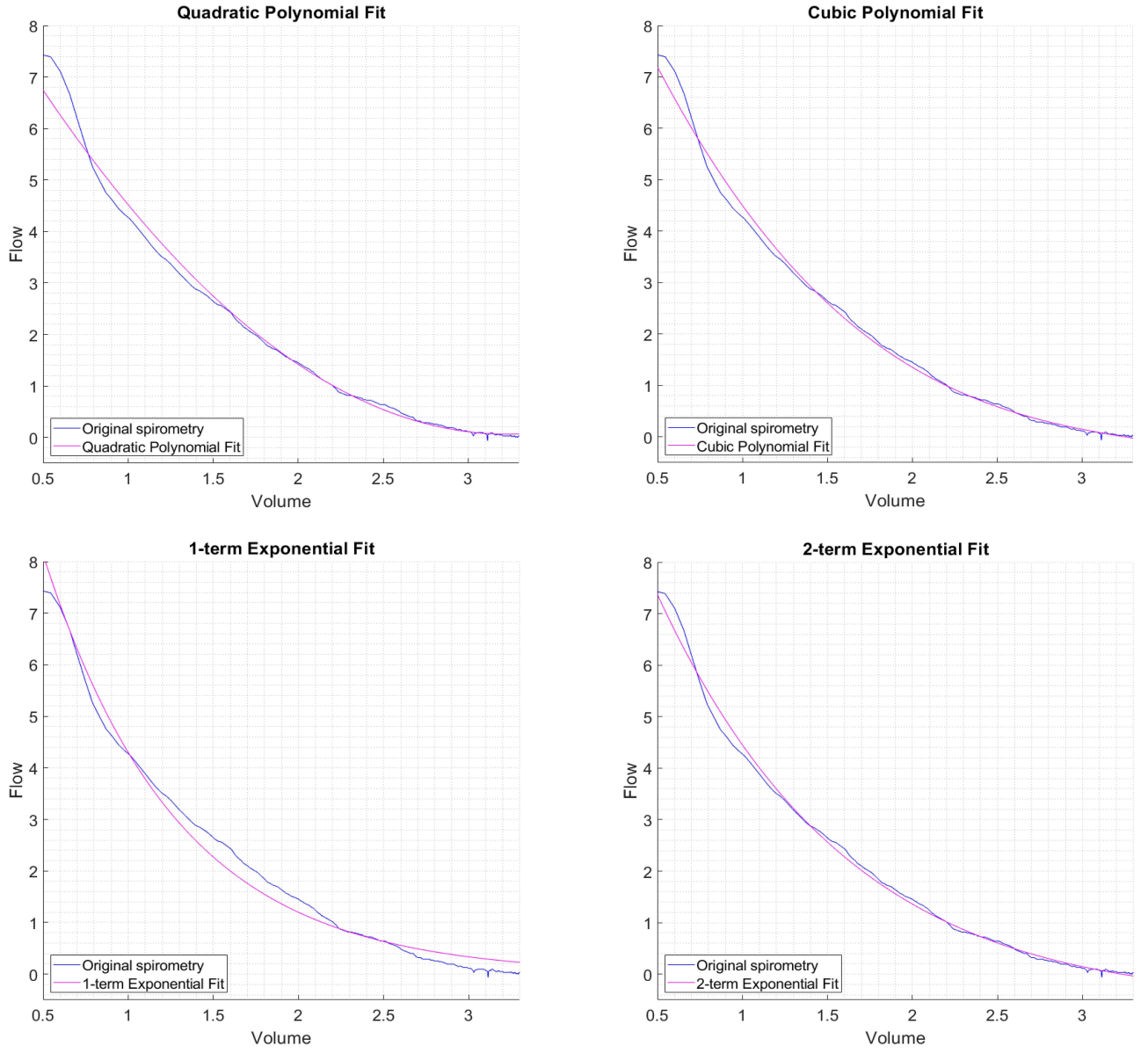

Figure 2 Example of curve fitting by quadratic polynomial, cubic polynomial, one-term and two term exponentials when applied to the spirometric curve.

distinguish between the two groups, we compared the behaviour of the coefficients between the COPD and No COPD groups, and we found that most of the coefficients had a low p-value, when evaluated individually. This result indicated that these features had the potential to correctly classify our subjects. Regardless of this result, we used all of them to feed the machine-learning techniques.

We chose five machine learning techniques to cover a wide range of possibilities. We used clustering techniques to explore the behaviour of our variables, but their diagnostic performance was not as good as that for supervised methods. Regarding the supervised techniques, we tried the simple decision tree because this algorithm could be easily understood and used in the clinical practice. Support Vector Machines were in the mid-range complexity, and we used it because it usually performs well in separating data in multi-dimensional spaces. Finally, the artificial neural network, the most complex technique, is difficult to interpret but it is the most used in the available literature, and it is the most powerful.[23] As expected, the artificial neural network had the best performance. Surprisingly, the best artificial network was the one fed by the coefficients of the quadratic polynomial, when the modelling is applied to the postbronchodilator curves.

Table 2 Curve Modelling Goodness of Fit

Model	SSE	R^2	adjR^2	RMSE
Quadratic polynomial – PreBD, *m*(IQR)	4.855 (2.342–10.309)	0.989 (0.978–0.994)	0.988 (0.978–0.994)	0.117 (0.082–0.17)
Quadratic polynomial – PostBD, *m*(IQR)	5.926 (2.641–11.972)	0.989 (0.978–0.994)	0.988 (0.978–0.994)	0.13 (0.085–0.19)
Cubic polynomial – PreBD, *m*(IQR)	2.169 (1.101–4.07)	0.995 (0.991–0.997)	0.995 (0.991–0.997)	0.08 (0.055–0.107)
Cubic polynomial – PostBD, *m*(IQR)	2.504 (1.263–4.71)	0.995 (0.992–0.997)	0.995 (0.992–0.997)	0.084 (0.06–0.119)
1-term exponential -PreBD, *m*(IQR)	20.214 (7.165–55.219)	0.955 (0.919–0.976)	0.955 (0.919–0.976)	0.238 (0.139–0.4)
1-term exponential – PostBD, *m*(IQR)	30.196 (11.283–72.742)	0.943 (0.907–0.971)	0.943 (0.907–0.971)	0.304 (0.179–0.47)
2-term exponential – PreBD, *m*(IQR)	4.77 (1.789–18.949)	0.989 (0.971–0.995)	0.989 (0.971–0.995)	0.117 (0.07–0.238)
2-term exponential – PostBD, *m*(IQR)	7.119 (2.492–23.785)	0.987 (0.969–0.994)	0.987 (0.968–0.994)	0.148 (0.084–0.266)

Abbreviations: *m*, median; IQR, interquartile range; PreBD, Prebronchodilator; PostBD, Postbronchodilator; SSE, Sum Squared Error; R^2, Coefficient of Determination; adjR^2, Adjusted Coefficient of Determination; RMSE, Root Mean Square Error.

Table 3 Coefficient Statistical Description

Variable	COPD	No COPD
Poly2-C_2-PreBD, *m*(IQR)	0.852 (0.609–1.295)	1.182 (0.751–1.769)*
Poly2-C_1-PreBD, *m*(IQR)	−4.421 (−5.627- −3.135)	−7.049 (−8.718- −5.65)*
Poly2-C_0-PreBD, *m*(IQR)	5.145 (3.569–7.677)	10.446(8.507–13.192)*
Poly2-C_2-PostBD, *m*(IQR)	0.887 (0.622–1.381)	1.106 (0.69–1.772)*
Poly2-C_1-PostBD, *m*(IQR)	−4.938 (−6.117- −3.674)	−7.274 (−9.071- −5.649)*
Poly2-C_0-PostBD, *m*(IQR)	6.375 (4.182–8.985)	11.447 (9.43–14.043)*
Poly3-C_3-PreBD, *m*(IQR)	−0.296 (−0.563- −0.085)	0.095 (−0.162–0.409)*
Poly3-C_2-PreBD, *m*(IQR)	2.454 (1.141–3.779)	0.345 (−1.618–2.398)*
Poly3-C_1-PreBD, *m*(IQR)	−6.121 (−8.451- −4.125)	−5.036 (−9.399- −1.594)*
Poly3-C_0-PreBD, *m*(IQR)	5.86 (3.997–8.771)	9.326 (7.158–12.111)*
Poly3-C_3-PostBD, *m*(IQR)	−0.183 (−0.456- −0.046)	0.241 (0.015–0.608)*
Poly3-C_2-PostBD, *m*(IQR)	1.828 (0.972–3.374)	−0.736 (−2.589–1.188)*
Poly3-C_1-PostBD, *m*(IQR)	−6.183 (−8.994- −4.309)	−3.862 (−7.541–0.251)*
Poly3-C_0-PostBD, *m*(IQR)	7.295 (4.75–9.518)	9.385 (6.87–12.218)*
Exp1-A-PreBD, *m*(IQR)	7.24 (4.715–11.981)	17.357 (13.322–22.683)*
Exp1-b-PreBD, *m*(IQR)	−1.605 (−2.019- −1.239)	−1.411 (−1.848- −1.082)*
Exp1-A-PostBD, *m*(IQR)	9.639 (6.091–14.017)	18.603 (14.934–25.287)*
Exp1-b-PostBD, *m*(IQR)	−1.478 (−1.95- −1.148)	−1.367 (−1.804- −1.057)*
Exp2-A-PreBD, *m*(IQR)	0 (−66.788–12.045)	6.341 (−109,949.792–160,430.449)*
Exp2-b-PreBD, *m*(IQR)	−1.108 (−2.279- −0.358)	−1.161 (−2.643- −0.486)*
Exp2-c-PreBD, *m*(IQR)	4.292 (−0.689–31.1)	0 (−160,420.368–109,965.042)
Exp2-d-PreBD, *m*(IQR)	−1.447 (−2.296- −0.878)	−1.272 (−2.643- −0.526)
Exp2-A-PostBD, *m*(IQR)	1.808 (−176.863–22,970.854)	8.593 (−103,208.119–185,791.764)
Exp2-b-PostBD, *m*(IQR)	−1.295 (−2.408- −0.557)	−0.846 (−2.411- −0.385)*
Exp2-c-PostBD, *m*(IQR)	3.052 (−22,965.667–204.025)	0 (−185,780.719–103,215.61)
Exp2-d-PostBD, *m*(IQR)	−1.522 (−2.408- −0.803)	−0.858 (−2.412- −0.406)*

Note: *p<0.05 Wilcoxon Rank Sum Test.
Abbreviations: COPD, Chronic Obstructive Pulmonary Disease; *m*, median; IQR, interquartile range; Poly2, Quadratic polynomial; Poly3, Cubic polynomial; Exp1, 1-term exponential; Exp2, 2-term exponential; PreBD, prebronchodilator; PostBD, postbronchodilator.

As shown in Table 7, our proposed system displays better diagnostic accuracy metrics than the current gold standard FEV1/FVC <0.7. Particularly, the specificity and the NPV show outstanding performance, which would make our classification system very valuable in a screening stage of the diagnostic process, such as in primary care.[26]

In a previous study, we proposed seven new measures of spirometry and evaluated their diagnostic accuracy.[27] Three of these measurements (slope at 50% of FVC, slope at PEF and slope at 75% of FVC) had an outstanding 95% accuracy

Table 4 Best Diagnostic Performance Results in Test Samples

Model	Accuracy	Sensitivity	Specificity	PPV	NPV
DT - Poly2 – PreBD	0.886	0.647	0.962	0.846	0.895
DT - Poly2 – PostBD	0.871	0.824	0.887	0.7	0.94
DT - Exp1 – PreBD	0.857	0.706	0.906	0.706	0.906
DT - Exp1 – All	0.857	0.706	0.906	0.706	0.906
SVM - Poly2 – PreBD	0.857	0.706	0.906	0.706	0.906
SVM - Poly2 – PostBD	0.857	0.706	0.906	0.706	0.906
SVM - Poly2 – All	0.886	0.706	0.943	0.8	0.909
SVM - Poly3 – PreBD	0.857	0.765	0.887	0.684	0.922
SVM - Poly3 – PostBD	0.857	0.647	0.925	0.733	0.891
SVM - Poly3 – All	0.886	0.765	0.925	0.765	0.925
SVM - Exp1 – PreBD	0.886	0.706	0.943	0.8	0.909
SVM - Exp1 – PostBD	0.857	0.647	0.925	0.733	0.891
SVM - Exp1 – All	0.886	0.706	0.943	0.8	0.909
ANN - Poly2 – PreBD	0.871	0.765	0.906	0.722	0.923
ANN - Poly2 – PostBD	**0.929**	**0.882**	**0.943**	**0.833**	**0.962**
ANN - Poly2 – All	0.914	0.824	0.943	0.824	0.943
ANN - Poly3 – PreBD	0.9	0.824	0.925	0.778	0.942
ANN - Poly3 – PostBD	0.9	0.824	0.925	0.778	0.942
ANN - Poly3 – All	0.871	0.882	0.868	0.682	0.958
ANN - Exp1 – PreBD	0.871	0.706	0.925	0.75	0.907
ANN - Exp1 – PostBD	0.857	0.706	0.906	0.706	0.906
ANN - Exp1 – All	0.871	0.706	0.925	0.75	0.907

Notes: The highlighted result is the Artificial Neural Network that was fed with the coefficients from the quadratic polynomial post-bronchodilator, which had the best diagnostic performance.

Abbreviations: PPV, Positive Predictive Value; NPV, Negative Predictive Value; DT, Decision Trees; SVM, Support Vector Machines; ANN, Artificial Neural Networks; Poly2, Quadratic polynomial; Poly3, Cubic polynomial; Exp1, 1-term exponential; Exp2, 2-term exponential; PreBD, Prebronchodilator; PostBD, Postbronchodilator.

Table 5 Confusion Matrix for the Gold Standard FEV1/FVC <0.7 on the Same 70 Patients That Were Used to Test the Machine-Learning Algorithm

		Expert Classification		
		COPD	No COPD	
FEV1/FVC<0.7	COPD	27	6	33
	No COPD	4	33	37
		31	39	70

Abbreviations: COPD, Chronic Obstructive Pulmonary Disease; FEV1, Forced Expiratory Volume at the first second; FVC, Forced Vital Capacity.

when compared to the FEV_1/FVC ratio. Other studies have also proposed new metrics for the curvature of the spirometry trace. For example, Das et al[9] found in a cohort of patients with COPD that the area under the forced expiratory flow-volume loop had good accuracy to detect severe hyperinflation. Mochizuki et al[10] proposed the obstruction index and found that it is very well correlated with emphysema in Computed Tomography (CT). In 2018, Bhatt et al[11] explored mathematical modelling using exponentials in the volume–time curve and found that Parameter D had a higher sensitivity than FEV_1/FVC for light obstruction when CT is used as reference.

Table 6 Confusion Matrix for the Testing Samples of the Artificial Neural Network Trained with the Coefficients from the Quadratic Polynomial After Applying Bronchodilator

		Expert Classification		
		COPD	No COPD	
ANN-Poly2-PostBD	COPD	15	3	18
	No COPD	2	50	52
		17	53	70

Abbreviations: COPD, Chronic Obstructive Pulmonary Disease; ANN, Artificial Neural Network; Poly2, Quadratic polynomial model; PostBD, Postbronchodilator.

Table 7 Diagnostic Accuracy Comparison Between FEV1/FVC <0.7 and the Artificial Neural Network Trained with the Coefficients from the Quadratic Polynomial After Applying Bronchodilator

Classification	Accuracy	Sensitivity	Specificity	PPV	NPV
FEV1/FVC<0.7	0.857	0.871	0.846	0.818	0.892
ANN Poly2 PostBD	0.929	0.882	0.943	0.833	0.962

Abbreviations: PPV, Positive Predictive Value; NPV, Negative Predictive Value; FEV1, Forced Expiratory Volume at the first second; FVC, Forced Vital Capacity; ANN, Artificial Neural Networks; Poly2, Quadratic polynomial; PostBD, Postbronchodilator.

On the other hand, machine learning techniques have been used for over a decade to facilitate the detection of respiratory diseases. Tang et al[15] evaluated the diagnostic performance of deep residual networks to automate COPD detection in low-dose CT and found that this technique can obtain very good accuracy and positive predictive values. Bodduluri et al[16] compared a deep neural network and a random forest classifier against traditional spirometric measures in their capabilities to differentiate among COPD structural phenotypes, showing the great potential of artificial intelligence for this diagnostic task. Finally, Ioachimescu et al[17] proposed the squared root of the area under the curve to feed an artificial neural network and it had a misclassification rate under 9% in patients with COPD diagnosis obtained through spirometry, plethysmography, and Helium dilution.

The main strength of this study is that we had access to raw spirometry data, which allowed us to work with actual spirometric data. Another of our strengths is that all studies were conducted in the same pulmonary lab, so all the records were acquired by professionals in respiratory therapy who had the same training and applied the same protocols.

One of the limitations of this study is that it was conducted with the data from a single centre, which may limit the extrapolation of the results. Finally, we did not have other more conclusive tests to diagnose COPD in the patients (such as Computed Tomography or Carbon monoxide diffusing capacity), which would guarantee a more certain diagnosis by the pneumologist.

Conclusion

This study explored the diagnostic performance of a combination of several curve modelling techniques and machine-learning techniques to diagnose COPD, based on spirometry alone. We found that an Artificial Neural Network fed with coefficients obtained from quadratic polynomial coefficients had 92.9% accuracy, 88.2% sensitivity and 94.3% specificity to detect COPD. The proposed system had better diagnostic accuracy performance than the current gold standard FEV1/FVC <0.7 in a sample of 70 patients. Hence, Artificial Neural Networks fed with quadratic and cubic polynomial coefficients have interesting potential for assisting the clinical diagnostic process, becoming an important aid in primary care to diagnose COPD in an early stage.

Patient and Public Involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Data Sharing Statement

Data are available upon reasonable request. The data that support the findings of this study are available on request.

Ethics Approval

This study involves human participants and was conducted according to the declaration of Helsinki, approved by the institutional ethics committee of the Clínica Universidad de La Sabana.

Acknowledgments

The study was carried out at the Universidad de La Sabana.

Author Contributions

All authors contributed to the study concept and design. Data acquisition was performed by AMF, ARBG, and DABR, and data analysis was performed by AMF, LFGC, MC, and ARBG. All authors contributed to the interpretation of the data. AMF, ETQ, and DABR drafted the work. LFGC, MC and ARBG critically revised every version of the article. All authors took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work. All authors have approved all changes and have accepted the final version for publication.

Funding

This work was supported by Universidad de La Sabana grant number MED-294-2020.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Global Initiative for Chronic Obstructive Lung Disease. POCKET GUIDE TO COPD DIAGNOSIS, MANAGEMENT, AND PREVENTION A guide for Health Care Professionals; 2023. Available from: https://goldcopd.org/wp-content/uploads/2023/03/POCKET-GUIDE-GOLD-2023-ver-1.2-17Feb2023_WMV.pdf. Accessed June 5, 2024.
2. Organización Mundial de la Salud. Enfermedad Pulmonar Obstructiva Crónica. 2023. Available from: https://www.who.int/es/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd). Accessed June 5, 2024.
3. Menezes AMB, Perez-Padilla R, Jardim JRB, et al. Chronic obstructive pulmonary disease in five Latin American cities (the PLATINO study): a prevalence study. *Lancet*. 2005;366(9500):1875–1881. doi:10.1016/S0140-6736(05)67632-5
4. Zou J, Sun T, Song X, et al. Distributions and trends of the global burden of COPD attributable to risk factors by SDI, age, and sex from 1990 to 2019: a systematic analysis of GBD 2019 data. *Respir Res*. 2022;23(1). doi:10.1186/s12931-022-02011-y
5. Global Initiative for Chronic Obstructive Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease - 2023 report; 2023; Available from: https://goldcopd.org/2023-gold-report-2/. Accessed June 5, 2024.
6. Celli BR. Update on the Management of COPD. *Chest*. 2008;133(6):1451–1462. doi:10.1378/chest.07-2061
7. Stanojevic S, Kaminsky DA, Miller MR, et al. ERS/ATS technical standard on interpretive strategies for routine lung function tests. *Eur Respir J*. 2022;60(1):1.
8. Casas Herrera A, De Oca MM, López Varela MV, Aguirre C, Schiavi E, Jardim JR. COPD underdiagnosis and misdiagnosis in a high-risk primary care population in four Latin American countries. A key to enhance disease diagnosis: the PUMA Study. *PLoS One*. 2016;11(4):e0152266. doi:10.1371/journal.pone.0152266
9. Das N, Topalovic M, Aerts JM, Janssens W. Area under the forced expiratory flow-volume loop in spirometry indicates severe hyperinflation in COPD patients. *Int J COPD*. 2019;14:409–418. doi:10.2147/COPD.S185931
10. Mochizuki F, Iijima H, Watanabe A, et al. The concavity of the maximal expiratory flow–volume curve reflects the extent of emphysema in obstructive lung diseases. *Sci Rep*. 2019;9(1):1–10. doi:10.1038/s41598-019-49591-2
11. Bhatt SP, Bhakta NR, Wilson CG, et al. New spirometry indices for detecting mild airflow obstruction. *Sci Rep*. 2018;8(1):1–8. doi:10.1038/s41598-018-35930-2
12. Oh A, Morris TA, Yoshii IT, Morris TA. Flow decay: a novel spirometric index to quantify dynamic airway resistance. *Respir Care*. 2017;62(7):928–935. doi:10.4187/respcare.04850

13. Li H, Liu C, Zhang Y, Xiao W. The concave shape of the forced expiratory flow-volume curve in 3 seconds is a practical surrogate of FEV1/FVC for the diagnosis of airway limitation in inadequate spirometry. *Respir Care*. 2017;62(3):363–369. doi:10.4187/respcare.05016
14. Johns DP, Das A, Toelle BG, et al. Improved spirometric detection of small airway narrowing: concavity in the expiratory flow–volume curve in people aged over 40 years. *Int J Chron Obstruct Pulmon Dis*. 2017;12:3567. doi:10.2147/COPD.S150280
15. Tang LYW, Coxson HO, Lam S, Leipsic J, Tam RC, Sin DD. Towards large-scale case-finding: training and validation of residual networks for detection of chronic obstructive pulmonary disease using low-dose CT. *Lancet Digit Health*. 2020;2(5):e259–67. doi:10.1016/S2589-7500(20)30064-9
16. Bodduluri S, Nakhmani A, Reinhardt JM, et al. Deep neural network analyses of spirometry for structural phenotyping of chronic obstructive pulmonary disease. *JCI Insight*. 2020;5(13). doi:10.1172/jci.insight.132781
17. Ioachimescu OC, Stoller JK. An alternative spirometric measurement: area under the expiratory flow–volume curve. *Ann Am Thorac Soc*. 2020;17(5):582–588. doi:10.1513/AnnalsATS.201908-613OC
18. Swaminathan S, Qirko K, Smith T, et al. A machine learning approach to triaging patients with chronic obstructive pulmonary disease. *PLoS One*. 2017;12(11):e0188532. doi:10.1371/journal.pone.0188532
19. Jafari S, Arabalibeik H, Agin K. Classification of Normal and Abnormal Respiration Pattern Using Flow Volume Curve and Neural Network. In: 5th International Symposium on Health Informatics and Bioinformatics IEEE. 2010:110–113.
20. Baemani J, Baemani MJ, Monadjemi A, Moallem P. Detection of respiratory abnormalities using artificial neural networks. *J Comput Sci*. 2008;4(8):663–667. doi:10.3844/jcssp.2008.663.667
21. Graham BL, Steenbruggen I, Barjaktarevic IZ, et al. Standardization of spirometry 2019 update an official American Thoracic Society and European Respiratory Society technical statement. *Am J Respir Crit Care Med*. 2019;200(8):E70–88. doi:10.1164/rccm.201908-1590ST
22. Machin D, Campbell MJ, Tan SB, Tan SH. *Sample Sizes for Clinical, Laboratory and Epidemiology Studies*. Wiley Online Library; 2018.
23. Hudson DL, Cohen ME. *Neural Networks and Artificial Intelligence for Biomedical Engineering*. New York, NY, USA: IEEE; 2000.
24. Berrar D. Cross-validation. In: *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*. Elsevier; 2018:542–545.
25. Šimundić AM. Measures of diagnostic accuracy: basic definitions. *EJIFCC*. 2009;19(4):203–211.
26. Maxim LD, Niebo R, Utell MJ. Screening tests: a review with examples. In: *Inhalation Toxicology*. Vol. 26. Informa Healthcare; 2014:811–828.
27. Moreno Giraldo AMM, Giraldo Cadavid LF, Botero Rosas D, et al. Comparison of new spirometry measures to diagnose COPD. *Respir Care*. 2022;2022:10191.