Taylor & Francis
Taylor & Francis Group

Expert Review of Respiratory Medicine

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ierx20

Demystification of artificial intelligence for respiratory clinicians managing patients with obstructive lung diseases

Joana Antão, Jeroen de Mast, Alda Marques, Frits M.E. Franssen, Martijn A. Spruit & Qichen Deng

To cite this article: Joana Antão, Jeroen de Mast, Alda Marques, Frits M.E. Franssen, Martijn A. Spruit & Qichen Deng (2023) Demystification of artificial intelligence for respiratory clinicians managing patients with obstructive lung diseases, Expert Review of Respiratory Medicine, 17:12, 1207-1219, DOI: 10.1080/17476348.2024.2302940

To link to this article: https://doi.org/10.1080/17476348.2024.2302940

EXPERT REVIEW OF RESPIRATORY MEDICINE
2023, VOL. 17, NO. 12, 1207–1219
https://doi.org/10.1080/17476348.2024.2302940

Taylor & Francis
Taylor & Francis Group

REVIEW

OPEN ACCESS Check for updates

Demystification of artificial intelligence for respiratory clinicians managing patients with obstructive lung diseases

Joana Antão[a,b,c,d], Jeroen de Mast[e], Alda Marques[a,b], Frits M.E. Franssen[c,d], Martijn A. Spruit[c,d] and Qichen Deng[c,d]

[a]Lab3R – Respiratory Research and Rehabilitation Laboratory, School of Health Sciences, University of Aveiro (ESSUA), Aveiro, Portugal; [b]iBiMED – Institute of Biomedicine, Department of Medical Sciences, University of Aveiro, Aveiro, Portugal; [c]Department of Research and Development, Ciro, Horn, The Netherlands; [d]Department of Respiratory Medicine, Maastricht University Medical Centre, NUTRIM School of Nutrition and Translational Research in Metabolism, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, the Netherlands; [e]Economics and Business, University of Amsterdam, Amsterdam, The Netherlands

ABSTRACT

Introduction: Asthma and chronic obstructive pulmonary disease (COPD) are leading causes of morbidity and mortality worldwide. Despite all available diagnostics and treatments, these conditions pose a significant individual, economic and social burden. Artificial intelligence (AI) promises to support clinical decision-making processes by optimizing diagnosis and treatment strategies of these heterogeneous and complex chronic respiratory diseases. Its capabilities extend to predicting exacerbation risk, disease progression and mortality, providing healthcare professionals with valuable insights for more effective care. Nevertheless, the knowledge gap between respiratory clinicians and data scientists remains a major constraint for wide application of AI and may hinder future progress. This narrative review aims to bridge this gap and encourage AI deployment by explaining its methodology and added value in asthma and COPD diagnosis and treatment.

Areas covered: This review offers an overview of the fundamental concepts of AI and machine learning, outlines the key steps in building a model, provides examples of their applicability in asthma and COPD care, and discusses barriers to their implementation.

Expert opinion: Machine learning can advance our understanding of asthma and COPD, enabling personalized therapy and better outcomes. Further research and validation are needed to ensure the development of clinically meaningful and generalizable models.

ARTICLE HISTORY
Received 13 July 2023
Accepted 4 January 2024

KEYWORDS
Artificial intelligence; asthma; chronic obstructive pulmonary disease; diagnosis; machine learning; management

1. Introduction

Artificial Intelligence (AI) has grown strikingly over the last decade, driven by recent technological advancements in computational power and increased acquisition and (real-time) availability of large volumes of different types of data. There is no consensus on AI definition [1]. For this review, the definition proposed in the University of Helsinki's Elements of AI course was adopted, which defines AI as systems that can execute specific tasks autonomously and adaptively. Broadly, AI includes machine learning (ML), rule-based expert systems and supporting technologies (Figure 1). AI is leading to a paradigm shift in clinical practice, optimizing processes and curtailing medical errors [2]. In respiratory medicine, it is primarily applied to optimize analyses of chest computed tomography scans and conventional chest radiographs, supporting the diagnosis of a wide range of health conditions, such as lung cancer [3].

Asthma and chronic obstructive pulmonary disease (COPD) are highly prevalent chronic airway diseases, placing a substantial burden on individuals, healthcare systems and societies [4,5]. AI offers great promise in addressing challenges associated with the diagnosis and management of these heterogeneous and complex diseases by i) analyzing simultaneously different types of data including demographics, lifestyle, patient-reported outcome measures, pulmonary and extrapulmonary features; and ii) finding linear and non-linear relationships and complex patterns. Consequently, it may aid in gaining insight into disease heterogeneity [6,7] and identifying patients at risk of exacerbations or premature death [8,9], allowing for timely treatment adjustments and improved health outcomes and costs.

Despite the mounting evidence showcasing AI advantages within respiratory medicine, its integration in clinical practice remains scarce. This is partly due to the unawareness of AI potential and clinicians' limited understanding of these techniques [10]. Bridging the knowledge gap may help build trust in AI, consequently boosting its deployment in respiratory medicine [10]. This narrative review aims to improve clinicians' understanding of AI and provide examples of its implementation in the diagnosis and management of chronic airway diseases.

CONTACT Joana Antão joanaantao@ua.pt Lab3R – Respiratory Research and Rehabilitation Laboratory, School of Health Sciences, University of Aveiro (ESSUA), Agras do Crasto - Campus Universitário de Santiago, Edifício 30, Aveiro 3810-193, Portugal

Supplemental data for this article can be accessed online at https://doi.org/10.1080/17476348.2024.2302940

Article highlights

- Artificial intelligence and machine learning can learn from big data, identify non-linear relationships, and uncover connections that healthcare professionals may overlook.
- Machine learning may improve the diagnosis and prognosis of asthma and chronic obstructive pulmonary disease by providing valuable support to physicians in clinical decision-making.
- A better understanding of machine learning by clinicians may overcome their resistance to the use of machine learning models and facilitate the widespread adoption of these techniques in healthcare.
- Future research should focus on conducting studies with larger samples and thoroughly validating machine learning models to ensure their generalizability and safety.
- Reinforcement learning and causal machine learning hold promise as future avenues for the management of chronic respiratory diseases.

2. Demystifying machine learning

ML, a subfield of AI, encompasses three main methods: supervised learning (deep and shallow), unsupervised learning, and reinforcement learning (Figure 1). For simplicity, this section introduces ML starting by covering linear regression.

A linear regression model is the relationship $y = f(x)$ between a dependent (or response) variable y, and ≥ 1 independent (or explanatory) variables x, as shown in (1):

$$\hat{y}_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \ldots + \beta_m x_{i,m} \quad (1)$$

Here, $x_1, \ldots, x_m$ denote the explanatory variables, which can be clinical characteristics (e.g. age or forced expiratory volume in the first second (FEV_1)). $\hat{y}_i$ denotes the response values predicted by the model (e.g. the predicted number of days until the next exacerbation-related hospitalization) as opposed to y_i, which are the true observed response values (e.g. the actual number of days until the exacerbation-related hospitalization).

The model's coefficients $\beta_0, \ldots, \beta_m$, are traditionally fitted by ordinary least squares which selects $\beta_0, \ldots, \beta_m$, such that the mean squared difference between observed and predicted/expected values is minimized (e.g. the smallest difference between the predicted and observed number of days until the next exacerbation-related hospitalization) (Figure S1).

Linear regression adjusted for confounders on observational data has been a common approach for hypothesis testing and causal explanation [11]. In the AI/ML era, emphasis has shifted to predictive models, which do not claim to test a causal hypothesis. Instead, they are correlational models, fitted on observational data, that claim to predict an outcome (e.g. the number of days until the next exacerbation-related hospitalization). Many clinical tasks are predictive in nature, such as identifying patients at increased risk of a poor outcome [12] or making a diagnosis from radiological images [13]. It is noteworthy that causality cannot be inferred from predictive models due to confounders. While randomized controlled trials are the gold standard for establishing causality, they are not always feasible [14]. With the increasing availability of data, there is a growing interest in estimating causal effects from observational data using ML (also known as causal ML) [14–17]. In contrast to traditional ML, causal ML combines data-driven methods with causal inference to determine the effect of a variable on an outcome while considering the complex interplay between all variables [15–17]. This usually requires domain knowledge of the relationships between variables involved to guide the modeling process and deduce causal effects [15–17]. Causal ML holds significant promise in healthcare, particularly for tasks that extend beyond prediction (e.g. deciding which intervention is likely to result in the best outcome [16]). In such cases, clinicians seek to understand what would happen to the outcome if different decisions were made [16]. In a recent application, causal ML has been used to estimate the treatment effects of dual therapy fluticasone furoate/vilanterol on mortality and exacerbation

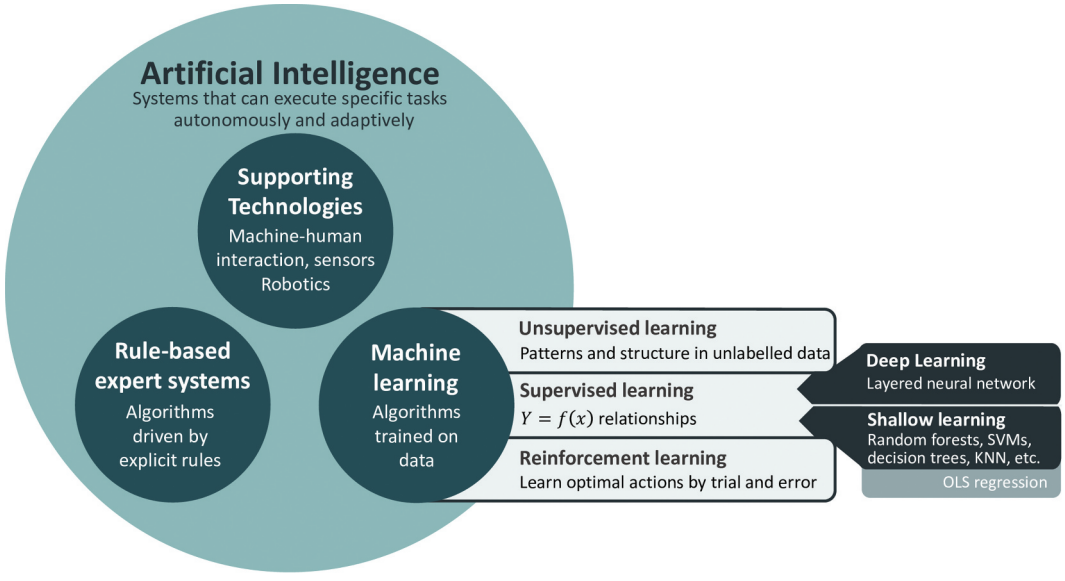

Figure 1. Artificial intelligence hierarchy. Artificial intelligence comprises a variety of methods, including supporting technologies, rule-based expert systems and machine learning. Machine learning can be further divided into supervised, unsupervised and reinforcement learning. KNN – K-nearest neighbors; OLS – ordinary least squares; SVM – support vector machine.

rate in people with COPD [18]. Healthcare professionals may, therefore, identify responses to different treatments prior to an intervention by incorporating causal ML in their practice.

2.1. Forms of machine learning

2.1.1. Supervised learning

Predictive modeling is the underlying concept of *supervised learning* and consists of models, fitted on data, that predict response values y (*labels* in ML) from independent variables x (*features* in ML) [19]. Linear regression is the simplest form of supervised learning. There are many types of algorithms, such as decision trees, support vector machines, and neural networks (Table S1). All these types of models are similar in form and function to the linear regression model (1), in that they are mathematical equations used to predict y values from x values. Supervised learning involves fitting a mathematical equation of the form $y = f(x)$ based on the training data, in which both the response y_i and the features x_i have been observed. The function of the model is to predict for new observations, where we observe the features x, and then use the equation to find a y value.

Problems in which the y variable is categorical are called *classification* (Figure 2(a)). For instance, support vector machine and artificial neural networks have been applied to discriminate asthmatics from healthy subjects (y_1=asthma or y_0=healthy), using lung sounds signals as predictors (features, $x's$) [20]. Conversely, if y is quantitative, then the application is called *regression* (Figure 2(b)). For example, predicting asthma lung function ($y = \text{FEV}_1\%$ predicted) from breathing and speech audio (features, $x's$) using linear regression, random forest, and support vector machine [21].

Deep learning takes model complexity to an extreme. A deep learning model is built around artificial neural networks, which are essentially mathematical equations of the form $y = f(x)$ [19]. In deep learning, artificial neural network models have multiple layers, where the output y of one layer is the input x for the next layer [22]. The resulting model is essentially still a mathematical equation of the form $y = f(x)$, but with many more terms than the model in (1). Deep neural networks have achieved high predictive performances in detecting asthma [23] and COPD [24] from sociodemographic variables, clinical data, biochemical results, lung function, bronchial challenge test [23], and lung sounds [24], respectively. Such algorithms are superior at handling many features, and therefore outperform shallow algorithms in applications, such as image and audio processing [25].

2.1.2. Unsupervised learning

Unsupervised learning is another form of ML, where x data (features) are available but no y values (labels) [19]. It consists of exploratory techniques, such as clustering (Figure 2(c)) and dimensionality reduction, which aim to find structure in data, or suggest a representation of data with fewer dimensions (i.e. fewer x variables), respectively. Table S1 provides a brief overview of the algorithms used for

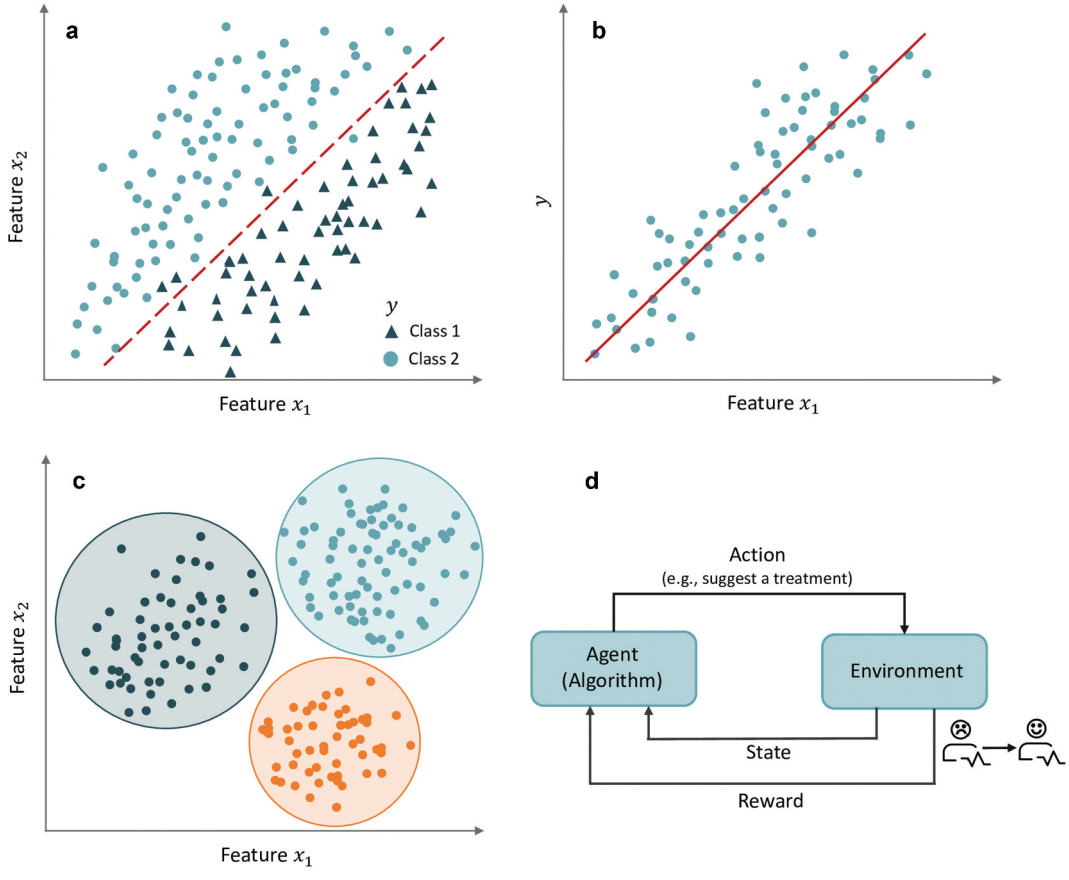

Figure 2. Graphical representation of (a) classification and (b) regression problems of supervised learning; (c) clustering analysis and (d) reinforcement learning.

these purposes. Clustering techniques have been of great value to better understand asthma and COPD heterogeneity. A hierarchical cluster analysis was conducted to explore COPD phenotypes based on 13 comorbidities (features, x, no label, y) [26].

2.1.3. Reinforcement learning

In reinforcement learning [27], the task is not to find a $y = f(x)$ relationship in the data, but to learn a suitable action in each set of possible scenarios. Conversely to supervised and unsupervised learning, reinforcement learning does not fit a model based on historical data, but instead, the procedure creates new data by interacting with the system in which it tries to learn suitable actions (Figure 2(d)). Reinforcement learning can be used for optimizing ventilation regimes in critically ill patients [28]. A ventilation strategy (tidal volume, positive end-respiratory pressure, and fraction of inspired oxygen) is proposed by the model based on patient's current characteristics (e.g. demographics, vitals, laboratory results, and medication). The algorithm learns the optimal ventilation regime by analyzing patients' health status and observing the result of its actions on patient's survival [28]. Reinforcement learning holds great potential in informing clinical decision-making due to its ability to address sequential decision-making problems, e.g. when treatment requires continuous adjustments considering changes in patient's health status [29]. Examples of successful research applications using reinforcement learning in healthcare practice [29] include the optimization of cancer treatment [29,30] and of multimorbidity management in patients with type 2 diabetes [31].

3. Rule-based expert systems

Rule-based algorithms translate expert knowledge, like the implicit rules used by physicians in diagnosis, into explicit if/then/else statements. Please note this is fundamentally different from ML; ML algorithms are mathematical relationships derived from data and not deduced from expert knowledge. Rule-based expert systems can support clinical decision making by providing recommendations or warnings from a set of rules based on patient's information. For example, an expert system has been developed by pulmonologists for supporting the diagnosis of obstructive lung disease in primary care [32]. Complex tasks like natural language processing and image recognition are challenging to handle with rule-based methods. As a result, ML has become the predominant approach, contributing to recent successes in AI.

4. Supporting technologies

AI also includes supporting technologies such as sensors, machine-human interaction, and knowledge representation, as well as integrated technologies in which ML is applied, such as robotics and autonomous cars. Examples of supporting technologies are the digital AI-powered stethoscope for remote monitoring of respiratory sounds [33], continuous remote monitoring of oxygen saturation levels using pulse oximetry [34] or physical activity sensors [35].

5. Phases of ML model development

ML models for medical purposes have advanced, but respiratory clinicians often lack understanding. This section describes the steps for building ML models.

5.1. Data cleaning

ML algorithms are data-driven methods capable of integrating extensive and multimodal data, i.e. different types of data including clinical and omics data, text, audio (e.g. lung sounds) and imaging (e.g. computed tomography scans). Healthcare data often contains missing values, inconsistencies, or errors, leading to inaccurate analysis and unreliable results that can jeopardize patient safety. It is imperative to ensure data is complete, accurate, and consistent before training a model. Data cleaning includes checking data accuracy, format, uniformity, de-duplication, handling of missing values and outliers [36].

5.2. Model selection

The selection of an ML algorithm is task dependent. If the aim is to predict an outcome such as a diagnosis (i.e. presence or absence of a disease) or prognosis (e.g. risk of death), supervised learning algorithms may be useful for these types of tasks. Conversely, if the goal is to uncover patterns in data where there is no predefined outcome variable, such as clustering individuals based on a set of characteristics, unsupervised learning methods should be employed. For tasks involving a sequence of decisions such as the management of chronic diseases with a sequential set of interventions to manage disease progression and severity, reinforcement learning becomes more appropriate. Other aspects should also be taken into consideration, such as number of features, model complexity, interpretability, and performance.

5.2.1. Number of features and model complexity

Including a large set of features may substantially increase model complexity and risk of overfitting. An ideal model has the right complexity to capture the $y = f(x)$ relationship, but not the random noise [19] (Figure S2A). Underfitted models fail to capture the $y = f(x)$ relationship and thereby produce less accurate predictions (Figure S2B). In contrast, overfitted models not only model the $y = f(x)$ relationship, but also capture sample-specific random noise [19] (Figure S2C), leading to a small error in the training set, but a large error in new data [19].

5.2.2. Predictive performance and interpretability

Predictive performance and interpretability are important aspects that influence the choice of a ML algorithm [36]. In healthcare, both accuracy and interpretability are required. Interpretable models, such as linear regression, enable easy understanding of how independent variables contribute to the response, but may show low predictive ability when input variables have non-linear relationships. In linear regression, a positive coefficient indicates that the mean of the dependent variable tends to increase as the

value of the independent variable increases, whereas a negative coefficient implies the opposite. More complex models, such as artificial neural networks, may yield higher predictive performances at the cost of lower interpretability [19], for which they are known as 'black box' algorithms [37].

5.3. Validation

Model validation is the process of evaluating an ML model on a test set to ensure its precision before using it in real-life applications [38]. There are two forms of validation: internal and external validation [39]. Internal validation is performed in individuals that have the same origin as the ones used for model training [39]. External validation consists of testing the model in a fully independent sample to determine its generalizability [39].

5.4. Performance metrics

Performance metrics are required to measure the performance of an ML model. Most regression metrics are based on the difference between the observed and the predicted values, such as mean absolute error and root mean squared error [40]. There is no determined threshold above which a model is considered appropriate, since both mean absolute error and root mean squared error are scale dependent. Therefore, as the error decreases, the model's performance improves.

In classification, most metrics are obtained from a confusion matrix, which is a cross-tabulation of the true and the predicted classes (Table S2) [40]. Several metrics can be calculated from the confusion matrix, such as sensitivity, specificity, and accuracy (Table S3) [40]. The model's performance improves as accuracy, sensitivity, and specificity converge toward one. A good discrimination model should have a high accuracy (>70%) [41] and a good trade-off between specificity and sensitivity (sum of specificity and sensitivity greater than 1.5) [42].

Another commonly reported metric is the area under the Receiver Operating Characteristic (ROC) Curve (AUC). The AUC can also be denoted as C-statistic. It is important to recognize that C-statistics involve distinct formulations in time-to-event analysis compared to the standard binary outcome. An AUC value of one indicates perfect classification, whereas an AUC value of 0.5 indicates that the model predicts no better than chance [19]. AUC scores above 0.7 are considered acceptable [43]. The AUC, however, can be misleading when classes are imbalanced [40]. In these situations, the area under the precision-recall curve is recommended [40].

6. Role of ML in asthma

Examples of articles implementing ML in asthma care are summarized in Table 1.

6.1. Diagnosis

Asthma diagnosis remains a major problem due to disease heterogeneity and non-specificity of the symptoms. In fact, respiratory symptoms vary over time and in intensity, which may lead to misattribution to other respiratory diseases, especially in the absence of lung function testing [44]. Therefore, diagnosis of asthma is often challenging, and attempts have been made to identify potential diagnostic markers using ML.

Sociodemographic variables, clinical data, spirometry parameters, biomechanical findings, and bronchial tests [23]; carbon dioxide waveforms [45]; respiratory sounds [20]; and Raman spectra from blood sera samples [46] have been considered in the development of ML-based diagnostic tools for asthma. Accuracies exceeded 90% using deep neural networks [23] and support vector machine [20,45,46].

6.2. Phenotypes

Asthma phenotyping is becoming increasingly important as it provides a foundation to understand disease etiology, heterogeneity and ultimately guide treatment [7]. Various clustering approaches, predominantly hierarchical, k-means, and k-medoids algorithms, have been employed to investigate asthma phenotypes in adults [7]. These approaches have considered a diverse range of variables, including sociodemographic, clinical, pathophysiological, lung function, behavioral, medication, and healthcare utilization data [7].

Based on readily available variables in clinical practice, recent research has identified three phenotypes using the k-medoids method [47]. The cluster solution was independently replicated in another sample [47]. Another recent study considered biomarkers from routine blood tests to explain asthma heterogeneity using the k-means algorithm. The three identified clusters had different risks of asthma exacerbations [48].

6.3. Management

Electronic health records have gained increased popularity as sources of valuable information and have already been considered for predicting asthma exacerbations [49–51]. Exacerbations, defined as the need for oral corticosteroids, an asthma-related emergency department visit, or hospitalization, were moderately predicted with a deep neural network (AUC = 0.70) [50]. Similarly, a boosting algorithm was used to predict non-severe exacerbations, emergency department visits or hospitalization [49]. Emergency department visits and hospitalizations were better predicted (AUC = 0.88 and AUC = 0.85, respectively) than non-severe exacerbations (AUC = 0.71) [49]. Similar accuracies were obtained when predicting emergency department visits or hospitalizations in asthma (AUC = 0.86) using another boosting algorithm [51]. Identified predictors for exacerbations were history of non-severe exacerbations requiring oral glucocorticoid bursts, severe asthma, age, number of hospital visits and number of systemic corticosteroids prescriptions [49,51].

Table 1. Summary of characteristics of the included studies applying machine learning for asthma diagnosis and management.

First author, year	Study population	Input features	Output	Algorithms	Validation	Performance
Tomita, 2019 [23]	566 subjects (367 people with asthma, 199 healthy controls)	Age, sex, cough, history wheezing, diurnal variation of symptoms, repeated symptoms, past history of allergy diseases, family history of allergy diseases, smoking status, current wheezing on auscultation, biochemical findings, spirometry parameters, bronchial challenge test	Asthma vs healthy controls	LR, SVM, DNN	10-fold stratified cross validation	DNN: Acc = 98%
Singh, 2018 [45]	73 subjects (43 people with asthma and 30 healthy controls)	Carbon dioxide waveforms	Asthma vs healthy controls	SVM, KNN, NB	LOOCV	SVM: Acc = 95%
Islam, 2018 [20]	60 subjects (30 people with asthma and 30 healthy controls)	Respiratory sounds	Asthma vs healthy controls	ANN, SVM	LOOCV	ANN (2- channel): Acc = 89% SVM (3-channel): Acc = 93%
Ullah, 2019 [46]	202 subjects (150 people with asthma, 52 healthy controls)	Raman spectral data	Asthma vs healthy controls	ANN, SVM, RF	10-fold cross validation	SVM: Acc = 94%, AUC = 0.90
Kisiel, 2020 [47]	1,291 people with asthma	Demographic and clinical characteristics	NA	k-medoids	Reproducibility in an independent sample	3 clusters
Oh, 2020 [48]	590 people with asthma	Blood biomarkers from routine blood tests	NA	k-means	Prospective validation with acute exacerbations	3 clusters
Xiang, 2020 [50]	31,433 people with asthma	Electronic health records	Exacerbation	LR, ANN, DNN	Train-test split; 5-fold cross validation	DNN: AUC = 0.70
Zein, 2021 [49]	60,302 people with asthma	Electronic health records	Non-severe exacerbations; emergency department visits and hospitalizations	LR, RF and LightGBM	Train-Test split	Non-severe exacerbations LightGBM: AUC = 0.71 Emergency department visits LightGBM: AUC = 0.88 Hospitalizations LightGBM: AUC = 0.85
Luo, 2020 [51]	334,564 entries (315,308 from 2005–2016 and 19,256 from 2017)	235 predictors	Emergency department visits or hospitalizations	40 algorithms	Cross-validation; Dataset from 2017 used for external validation	XGBoost: AUC = 0.86
Hafke-Dys, 2021 [33]	899 people with asthma (19,166 recordings)	Indexes from a digital AI-powered stethoscope	Abnormal or normal cough	DNN	Train-Test split	AUC = 0.94
Khasha, 2019 [52]	96 people with asthma	Demographic characteristics, spirometry parameters, symptoms, medical history and environmental factors	Asthma control level (very poorly controlled, not well controlled, well controlled)	LR, SVM, RF, XGBoost, KNN, DT, NB, rule-based model from clinicians' knowledge	5-fold cross-validation	Machine learning+rule-based model Acc = 91.7%
Halner, 2021 [53]	81 subjects (59 people with asthma; 22 people with COPD)	Demographics, medical history, medication, VAS, MRC dyspnea scale, HADS, EuroQol 5D, vital signs, lab results	Systemic Corticosteroids prescription and treatment failure	RF	LOOCV	Systemic corticosteroids Prescription AUC = 0.69 Treatment Failure AUC = 0.81

Acc: Accuracy; ANN: Artificial Neural Network; AUC: Area under the Receiver Operating Characteristic Curve; COPD: chronic obstructive pulmonary disease; DT: Decision tree; DNN: Deep Neural Network; XGBoost: extreme gradient boosting; HADS: Hospital Anxiety and Depression Scale; KNN: k-nearest neighbor; LOOCV: Leave-One-Out Cross-Validation; LightGBM: light gradient-boosting machine; LR: Logistic regression; MRC: Medical Research Council; NA: Not applicable; NB: Naïve Bayes; RF: Random Forest; SVM: Support Vector Machine; VAS: Visual Analogue Scale.

Monitoring asthma control levels and symptoms can guide clinical decision-making. An artificial neural network model was proposed to remotely monitor asthma symptoms using indexes from a digital AI-powered stethoscope [33]. Included were individuals of all age groups, both asthmatics and non-asthmatics, with and without adventitious respiratory sounds [33]. The model that best distinguished abnormal from normal respiratory sounds in people with asthma reached an AUC of 0.94 using intensity scores of wheezes and rhonchi [33]. Recently, a novel approach combined a rule-based expert model and ML to predict asthma control level from demographics, clinical characteristics, lung function and environmental factors, yielding an excellent predictive power (accuracy = 91.7%) [52].

ML has also shown potential in identifying patients who may be responsive to standard care or who may require more personalized treatment. A random forest algorithm predicted treatment failure after an exacerbation and the prescription of systemic corticosteroids during exacerbation in individuals with asthma and COPD [53]. Patients who were readmitted, required treatment adjustment, or died over the 30-day follow-up period were deemed unresponsive [53]. The model achieved a good performance for treatment failure (AUC = 0.81), with the scores of the visual analogue scale for breathlessness and sputum purulence being the most predictive [53]. A satisfactory performance (AUC = 0.69) was found for the prescription of systemic corticosteroids using only as input the presence of wheezing and the percentage of blood eosinophils [53].

7. Role of ML in COPD

Examples of the applicability of ML in COPD care are described in Table 2.

7.1. Diagnosis

Several attempts have been made to diagnose COPD using different types of data. Data from cardiopulmonary exercise tests [54], chest computed tomography scans [55,56], respiratory sounds [24] coupled with spirometry [57], volatile organic compounds [58,59], electronic health records [60] have been used to distinguish COPD from controls. Accuracies varied between 76.7% to 100% using support vector machine [54,57,58] or boosting algorithms [59]. The model developed based on real-world data also reached an excellent predictive performance (precision-recall AUC = 0.93) [60]. Deep learning applied to lung sounds (sensitivity = 0.93, specificity = 0.93) [24] and chest computed tomography scans (AUC >0.85) [55,56] also showed high discriminatory power in the detection of COPD.

The forced oscillation technique parameters have been used to build classifiers that discriminate between different levels of airflow limitation in people with COPD [61]. K-nearest neighbors and random forest yielded excellent performances (AUC ≥0.90) in control versus all people with COPD and control versus patients with moderate, severe, and very severe airflow limitation [61].

7.2. Phenotypes

Over the last decade, considerable research was performed to identify clinically relevant COPD phenotypes using unsupervised learning, mainly through hierarchical and k-means algorithms [6].

COPD heterogeneity has also been explored using k-means [62], k-medoids [63] and hierarchical [64] algorithms considering multidimensional data [62] as well as clinically relevant and easily accessible variables [63,64]. Clusters were either prospectively validated with clinically relevant outcomes [62,64] or assessed for stability over time [63]. Other studies demonstrated that people with COPD have distinct comorbidity [26] and lung function profiles [65] using self-organizing maps followed by hierarchical clustering. There are, however, several studies reporting different cluster solutions, which reflect the use of different sample sizes and characteristics, choice of input variables and clustering algorithms [6].

7.3. Management

Currently, risk classification of COPD is mainly based on the history of previous exacerbations and hospitalizations, which is often based on patient recall and characterized by underreporting of events [5]. Early identification of at-risk patients may allow timely treatment adjustment, prevent disease progression, and reduce the burden on the healthcare system. Recent research has applied ML to identify risk factors for hospital readmission in people with COPD [66]. Top predictors were hospitalization in the previous two years, older age, being male, number of comorbidities, and longer length of hospital stay [66].

Exacerbation prediction based on remote monitoring of respiratory sounds has also been reported [67]. A support vector machine model was able to predict the onset of exacerbation, on average, 5 ± 1.9 days before in 75.8% of the cases [67].

ML has also been applied to identify physically inactive people with COPD who could benefit from physical activity promotion interventions [68]. Patients were divided into two categories based on their daily walking duration and intensity level: extremely inactive or overactive. The random forest algorithm, which was trained to distinguish extremely inactive from overactive using nonphysical activity related data, yielded an overall AUC of 0.84 [68].

Prediction models have also been developed to identify individuals at risk of mortality using decision tree algorithm [12], Cox regression [69], and deep neural networks [55]. The tree-based model developed based on age, spirometry parameters, dyspnea, physical activity, and number of hospital admissions in the previous 2 years reached comparable performances when compared to established mortality prediction models (AUC≈0.7) [12]. Similarly, a Cox regression model built using imaging, spirometry, and clinical data to predict all-cause mortality in people with COPD [69], outperformed body-mass index, airflow obstruction, dyspnea, exercise capacity (BODE) index, modifications of BODE index and age, dyspnea, and airflow obstruction index [69]. The top predictors were the 6-minute walk distance, FEV_1%-predicted, age, and pulmonary

Table 2. Summary of characteristics of the included studies applying machine learning for COPD diagnosis and management.

First author, year	Study population	Input features	Output	Algorithms	Validation	Performance
Inbar, 2021 [54]	234 subjects (73 people with chronic heart failure, 75 people with COPD and 86 healthy controls)	CEPT data	Chronic heart failure vs COPD vs healthy controls	SVM	Train-test split; cross-validation	Acc = 99%
Srivastava, 2021 [24]	126 subjects	Respiratory sounds	COPD vs healthy controls	DNN	Train-test split; 10-fold cross-validation	Sn = 0.93; Sp = 0.93
Haider, 2019 [57]	55 subjects (30 people with COPD and 25 healthy controls)	Respiratory sounds and spirometry parameters	COPD vs healthy controls	SVM, KNN, LR, DT and discriminant classifiers	Train-test split; 5-fold and 10-fold cross-validation	Lung sounds SVM: Acc = 83.6% Lung sounds + spirometry SVM: Acc = 100%
Amaral, 2015 [61]	168 subjects (126 people with COPD and 42 healthy controls)	Forced oscillation technique parameters	COPD different levels of obstruction vs healthy controls	SVM (Linear and radial basis kernel), KNN, RF	10-fold cross-validation	Control versus all people with COPD and control versus moderate, severe, and very severe people with COPD separately AUC >0.90
Berkel, 2010 [58]	79 subjects (50 people with COPD, 29 healthy controls)	Volatile organic compounds identified by gas chromatography-mass spectroscopy	COPD vs healthy controls	SVM	Cross-validation, external validation (16 people with COPD and 16 controls)	Cross-validation Acc = 92% Test set Acc = 91%
V.A B, 2021 [59]	199 subjects (55 people with COPD, 51 people with lung cancer and 93 healthy controls)	Volatile organic compounds identified by an electronic nose device	COPD vs healthy controls, lung cancer vs healthy controls	XGBoost, AdaBoost, RF	Train-test split; 3-fold, 5-fold and 10-fold cross-validation	COPD vs Healthy controls XGBoost: Acc = 77%, AUC = 0.76
González, 2018 [55]	7,983 COPD smokers (8,983 from COPDGene study and 1,672 from ECLIPSE study)	Chest computed tomography imaging	COPD vs controls, 3-year all-cause mortality	DNN	Train-test split of the sample from COPDGene study, ECLIPSE sample used for external validation	COPD vs. controls AUC = 0.856 Mortality C-statistic = 0.60
Tang, 2020 [56]	4,784 COPD former and current smokers (2,589 from PanCan study and 2,195 from ECLIPSE)	Chest computed tomography imaging	COPD vs healthy controls	DNN	3-fold cross-validation, ECLIPSE sample used for external validation	AUC = 0.886
Mariani, 2021 [60]	19,077 people with asthma or COPD (2007–2011)	Electronic Health Records	COPD vs asthma vs ACOS	SVM (linear and radial basis function), KNN, RF	Train-test split; 10 to 30-fold cross-validation	pre-rec AUC Asthma = 0.9, ACOS = 0.34, COPD = 0.93, unclear = 0.25
Garcia-Aymerich, 2011 [62]	342 people with COPD hospitalized with exacerbation	Respiratory symptom and quality of life, lung function, chest x-ray morphometry, systematic inflammation, nutritional status, muscle strength, exercise capacity, comorbidities, lung density, airway morphology, bronchial colonization and inflammation data	NA	k-means	Prospective validation with hospital admission and all-cause mortality	3 clusters
Burgel, 2017 [64]	2,409 people with COPD	Age, BMI, spirometry parameters, mMRC dyspnea scale, number of previous exacerbations, presence of comorbidities	5 clusters	Hierarchical algorithm, DT	Train-test split; Prospective validation with all-cause mortality	5 clusters DT Acc = 77%
Marques, 2022 [63]	352 people with COPD	Sociodemographic and clinical data, lung function, symptoms, disease impact, quality of life, lower-limb muscle strength and functional status	4 clusters	k-medoids, DT	Train-test split; stability over time	4 clusters DT Acc = 71.7%

(*Continued*)

Table 2. (Continued).

First author, year	Study population	Input features	Output	Algorithms	Validation	Performance
Vanfleteren, 2013 [26]	213 people with COPD	13 comorbidities (chronic kidney disease, anemia, hypertension, obesity, underweight, muscle wasting, hyperglycemia, dyslipidemia, osteoporosis, anxiety and depression, atherosclerosis, myocardial infarction)	NA	Self-organizing maps and hierarchical algorithm	NA	5 clusters
Augustin, 2018 [65]	518 people with COPD	Lung function attributes	NA	Self-organizing maps and hierarchical algorithm	NA	7 clusters
Cavailles, 2020 [66]	143,006 people with COPD hospitalized for an acute exacerbation	Sociodemographic and clinical variables	Rehospitalization for acute exacerbation	DT	Train-test split	NA
Fernandez-Granero, 2015 [67]	15 people with COPD	Respiratory sounds	Acute exacerbations	SVM	10-fold cross validation	Sn = 0.74, Sp = 0.98
Aguilaniu, 2021 [68]	1,409 people with COPD	Clinicopathological data	Extreme inactivity	RF	Train-test split	AUC = 0.84
Moll, 2020 [69]	3,900 people with COPD (2,632 from COPDGene study and 1,268 from ECLIPSE)	Clinical, spirometry and imaging features	All-cause mortality	Cox regression with RF	Train-test split of the sample from COPDGene study, ECLIPSE sample used for external validation	Cox regression C-statistic = 0.702 BODE C-statistic = 0.66 BODE and exacerbations C-statistic = 0.67 updated BODE C-statistic = 0.65 ADO C-statistic = 0.65
Esteban, 2011 [12]	611 people with COPD	Sociodemographic data, spirometry parameters, symptoms, physical activity and number of hospital admissions in the previous 2 years	5-year mortality	DT	External validation (348 people with COPD)	DT AUC = 0.74 BODE AUC = 0.72 HADO AUC = 0.70 ADO AUC = 0.74
Luo, 2020 [70]	263,015 people with COPD (780,295 hospital admissions)	Medical insurance data (demographic, clinical and costs information)	High-cost patients	LR, XGBoost, RF	Train-test split and external validation (2 datasets)	Train-test XGBoost: AUC = 0.80 Test set 1 XGBoost: AUC = 0.78 Test set 2 XGBoost: AUC = 0.76

Acc: Accuracy; AdaBoost: adaptive boosting; ADO: age, dyspnea and airflow obstruction; pre-rec AUC: Area under precision-recall curve; AUC: Area under the Receiver Operating Characteristic Curve; BMI: Body mass index; BODE: body mass index, airflow obstruction, dyspnea, and exercise capacity; CEPT: Cardiopulmonary Exercise Test; COPD: chronic obstructive pulmonary disease; DT: Decision tree; DNN: Deep Neural network; XGBoost: extreme gradient boosting; HADO: health, activity, dyspnea, obstruction; KNN: k-nearest neighbors; LR: Logistic regression; mMRC: modified Medical Research Council; NA: Not applicable; RF: Random Forest; Sn: Sensitivity; Sp: Specificity; SVM: Support Vector Machine.

artery-to-aorta ratio [69]. In contrast, a deep learning model exhibited only moderate predictive capability for the risk of death, using computed tomography imaging data (C-statistic = 0.6) [55].

Moreover, ML has been used to identify high-cost patients based on medical insurance data [70]. A good predictive performance was reached with a boosting algorithm (AUC = 0.80). Relevant predictors were cost-related variables, age, region, gender, type of insurance, number of comorbidities, emphysema, hypertension, heart disease, and Charlson Comorbidity Index scores [70].

8. Implementation challenges

AI deployment into outpatient or bedside care for asthma and COPD has been hindered by the knowledge gap but also by other hurdles. Studies often lack external validation, which may lead to overoptimistic predictive performances [71]. Cluster validity assessment is crucial to confirm reproducibility of the proposed classification and their relevance in guiding asthma and COPD care. Studies also frequently included small sample sizes, which hinders the model's generalizability. Hence, studies with larger samples, as well as external validation of findings are crucial steps prior to ML implementation within clinical workflow [71]. Additionally, it is also imperative to compare the predictive ability across models and conduct prospective clinical trials to assess efficacy at improving patient care [72].

ML also raises concerns regarding data protection [36]. Data collection and processing must comply with ethical guidelines and data privacy legislation throughout model development and implementation [36].

Another major issue arises when black-box algorithms are employed. The lack of interpretability delays ML acceptance, as clinicians are unable to double-check for errors or biases and provide an explanation to the patient [36,72].

Accountability is also a source of ongoing debate in AI implementation in healthcare, as poor decision-making may jeopardize patient safety [72]. Currently, no broad consensus exists regarding responsibility for the model's decisions since healthcare professionals, data scientists, and software developers lack complete control [72].

9. Conclusion

This review covers fundamental concepts of AI/ML and main steps of ML model development, providing a simple guide for clinicians to improve their understanding of AI/ML, create trust, and accelerate ML deployment in clinical practice. This paper has also provided examples of AI/ML for diagnosis, phenotyping, and management of people with asthma and COPD. AI may aid in making informed decisions about diagnosis and management by incorporating a large number of data and drawing connections that healthcare professionals may overlook. AI has potential to enable a fast and accurate diagnosis and support early identification of at-risk patients. It also provides valuable information necessary for treatment adjustments and personalized medicine, preventing worsening of patient's health, saving time and resources.

However, low-quality evidence, data privacy, interpretability, and accountability need to be addressed for AI to become a reality in healthcare. Research with larger sample sizes and a thorough validation of findings are crucial to ensure unbiased, generalizable, and clinically relevant ML models.

10. Expert opinion

This review introduces ML in a user-friendly language and gives examples of its application in the context of respiratory medicine. It describes how ML can support the diagnosis and management of chronic respiratory diseases such as asthma and COPD, which are characterized by considerable heterogeneity and complexity. It also serves as a guide for clinicians to improve their understanding and build trust in the use of ML in healthcare. Nevertheless, clinicians need to be aware of some methodological issues to be considered in future studies with ML. Adequate sample size, high quality data and thorough validation are critical requirements for developing robust ML models. A sufficient sample size and representativeness of the target population are crucial to ensure generalizability and reduce the risk of overfitting and bias toward a particular group. Electronic health records have become one of the most important sources of data in clinical research. However, data collected in clinical practice are often unstructured, non-standardized and incomplete, which can significantly affect the performance of the model if not properly addressed. It is necessary to report errors that occur and describe the pre-processing steps used to correct them. Establishing the performance of a model should include not only internal, but also external validation. Many models have demonstrated excellent predictive performance during development phase, only to fail when tested on an independent sample. Therefore, extensive validation is needed to ensure generalizability and avoid potentially harmful models for patients. Interpretability should also be highlighted as an important aspect in the development of ML models. Simpler models allow for better transparency of the decision-making process, which can help detect and reduce algorithmic bias and clarify some of the ethical issues associated with the use of AI (e.g. accountability), while promoting trust among healthcare professionals and patients. Post-hoc explainability methods have been developed to facilitate understanding of the reasons for the predictions of black-box models. However, there is still room for improvement [73,74]. In addition, the (dis)advantages of incorporating ML models into standard care need to be carefully weighed from both a patient-centered and an economic perspective [36], which is only possible if data scientists work closely with healthcare professionals and other stakeholders to identify and resolve important issues early on. Such efforts can lead to models that are clinically useful and operationally feasible. Finally, it is of utmost importance to follow the guidelines for development (e.g. Cross-Industry Standard Process for Data Mining – CRISP-DM [75,76]) and reporting of ML projects (e.g. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis – TRIPOD [77]).

The potential impact of reinforcement learning in the management of chronic diseases is also noteworthy. In reinforcement learning, data is constantly fed into the system and

incorporated into the model through continuous updates. Unlike static models – i.e. the model is trained exactly once – the goal is to choose an action at each point in time that maximizes long-term reward through constant feedback from the environment. Reinforcement learning can become a powerful tool for improving chronic respiratory disease management and represents a promising avenue for future advances in the field. Moreover, with its ability to identify treatment responses before intervention, causal ML can support clinicians in selecting the most effective treatment for each patient, and therefore may also represent a promising avenue to advance personalized medicine in chronic respiratory diseases.

Abbreviations

AI	Artificial Intelligence
AUC	Area under the Receiver Operating Characteristic Curve
BODE	Body mass index, airflow obstruction, dyspnea, exercise capacity
COPD	Chronic obstructive pulmonary disease
FEV_1	Forced expiratory volume in one second
ML	Machine learning
ROC	Receiver Operating Characteristic

Funding

Part of the work of J Antão and Q Deng is financially supported by AstraZeneca, Boehringer Ingelheim, Chiesi and TEVA.

Declaration of interest

FME Franssen has obtained research grants from AstraZeneca, outside the scope of the current study. FME Franssen has obtained consultancy fees from MSD for advisory boards outside the scope of the current study. FME Franssen has received speakers' fees by AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, Chiesi and Novartis. MAS has obtained research grants from Netherlands Lung Foundation and Stichting Astma Bestrijding, outside the scope of the current study. MA Spruit has obtained research grants from AstraZeneca, TEVA, Chiesi and Boehringer Ingelheim for the current study. MA Spruit has obtained consultancy fees from AstraZeneca and Boehringer Ingelheim for advisory boards outside the scope of the current study. All research grants and consultancy fees were paid to Ciro. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Reviewer disclosures

A reviewer on this manuscript received an honorarium from Expert Review of Respiratory Medicine for their review work. Peer reviewers on this manuscript have no other relevant financial relationships or otherwise to disclose.

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Prins C, Sheikh H, Schrijvers E, et al. Mission AI: The New System Technology. The Hague: The Netherlands Scientific Council for Government Policy (WRR); 2021.
2. Hazarika I. Artificial intelligence: opportunities and implications for the health workforce. Int Health. 2020 Jul 1;12(4):241–245. doi: 10.1093/inthealth/ihaa007
3. Diagnostic Image Analysis Group. AI for radiology: an implementation guide. Available from: https://grand-challenge.org/aiforradiology/?subspeciality=Chest&modality=All&search
4. Global Initiative for Asthma (GINA). 2022 GINA report - global strategy for asthma management and prevention. 2022
5. Global Initiative for Chronic Obstructive Lung Disease (GOLD). GOLD report - global strategy for the diagnosis, management and prevention of COPD. 2022. 2022.
6. Nikolaou V, Massaro S, Fakhimi M, et al. COPD phenotypes and machine learning cluster analysis: a systematic review and future research agenda. Respir Med. 2020 Sep;171:106093. doi: 10.1016/j.rmed.2020.106093
• **Systematic review of the applicability of unsupervised learning in profiling COPD patients.**
7. Cunha F, Amaral R, Jacinto T, et al. A systematic review of asthma phenotypes derived by data-driven methods. Diagnostics. 2021 Apr 2;11(4):644. doi: 10.3390/diagnostics11040644
• **Systematic review of the applicability of unsupervised learning in profiling asthma patients.**
8. Exarchos KP, Beltsiou M, Votti CA, et al. Artificial intelligence techniques in asthma: a systematic review and critical appraisal of the existing literature. Eur Respir J. 2020 Sep;56(3):2000521. doi: 10.1183/13993003.00521-2020
• **Systematic review of the applicability of artificial intelligence in asthma diagnosis and management.**
9. Exarchos KP, Aggelopoulou A, Oikonomou A, et al. Review of artificial intelligence techniques in chronic obstructive lung disease. IEEE J Biomed Health Inform. 2022 Dec 16;26(5):2331–2338. doi: 10.1109/JBHI.2021.3135838
• **Review of the applicability of artificial intelligence in COPD diagnosis and management.**
10. Pucchio A, Eisenhauer EA, Moraes FY. Medical students need artificial intelligence and machine learning training. Nat Biotechnol. 2021 Mar 01;39(3):388–389.
11. Shmueli G. To explain or to predict? Stat Sci. 2010;25(3):289–310. doi: 10.1214/10-STS330
12. Esteban C, Arostegui I, Moraza J, et al. Development of a decision tree to assess the severity and prognosis of stable COPD. Eur Respir J. 2011;38(6):1294–300. doi: 10.1183/09031936.00189010
13. Chen S, Suzuki K, MacMahon H. Development and evaluation of a computer-aided diagnostic scheme for lung nodule detection in chest radiographs by means of two-stage nodule enhancement with support vector classification. Med Phys. 2011 Apr;38(4):1844–58. doi: 10.1118/1.3561504
14. Olier I, Zhan Y, Liang X, et al. Causal inference and observational data. BMC Med Res Methodol. 2023 Oct 11;23(1):227. doi: 10.1186/s12874-023-02058-5
15. Bica I, Alaa AM, Lambert C, et al. From real-world patient data to individualized treatment effects using machine learning: Current and future methods to address underlying challenges. Clin Pharmacol Ther. 2021 Jan;109(1):87–100.
16. Prosperi M, Guo Y, Sperrin M, et al. Causal inference and counterfactual prediction in machine learning for actionable healthcare. Nat Mach Intell. 2020 Jun 01;2(7):369–375. doi: 10.1038/s42256-020-0197-y
•• **Importance of causal inference in personalized medicine.**
17. Sanchez P, Voisey JP, Xia T, et al. Causal machine learning for healthcare and precision medicine. R Soc Open Sci. 2022 Aug;9(8):220638.
18. Verstraete K, Gyselinck I, Huts H, et al. Estimating individual treatment effects on COPD exacerbations by causal machine learning on randomised controlled trials. Thorax. 2023 Oct;78(10):983–989.
19. WD JG, Hastie T, Tibshirani R. An introduction to statistical learning. (New York): Springer; 2013.
20. Islam MA, Bandyopadhyaya I, Bhattacharyya P, et al. Multichannel lung sound analysis for asthma detection. Comput Methods Programs Biomed. 2018 Jun;159:111–123.
21. Alam MZ, Simonetti A, Brillantino R, et al. Predicting pulmonary function from the analysis of voice: a machine learning approach. Front Digit Health. 2022;4:750226. doi: 10.3389/fdgth.2022.750226

22. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015 May 01;521(7553):436–444.
23. Tomita K, Nagao R, Touge H, et al. Deep learning facilitates the diagnosis of adult asthma. Allergol Int. 2019 Oct;68(4):456–461.
24. Srivastava A, Jain S, Miranda R, et al. Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease. PeerJ Comput Sci. 2021;7:e369. doi: 10.7717/peerj-cs.369
25. Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. Electron Markets. 2021 Sep 01;31(3):685–695.
26. Vanfleteren LE, Spruit MA, Groenen M, et al. Clusters of comorbidities based on validated objective measurements and systemic inflammation in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2013 Apr 1;187(7):728–735. doi: 10.1164/rccm.201209-1665OC
27. Sutton RS, Barto AG. Reinforcement learning: an introduction. 2nd ed. Cambridge: The MIT Press; 2018.
28. Peine A, Hallawa A, Bickenbach J, et al. Development and validation of a reinforcement learning algorithm to dynamically optimize mechanical ventilation in critical care. NPJ Digit Med. 2021 Feb 19;4(1):32. doi: 10.1038/s41746-021-00388-6
29. Yu C, Liu J, Nemati S, et al. Reinforcement learning in healthcare: a survey. ACM Comput Surv. 2023;55(1):1–36. doi: 10.1145/3477600
•• **Overview of reinforcement learning applications in healthcare.**
30. Padmanabhan R, Meskin N, Haddad WM. Reinforcement learning-based control of drug dosing for cancer chemotherapy treatment. Math Biosci. 2017 Nov;293:11–20. doi: 10.1016/j.mbs.2017.08.004
31. Zheng H, Ryzhov IO, Xie W, et al. Personalized multimorbidity management for patients with type 2 diabetes using reinforcement learning of electronic health records. Drugs. 2021 Mar;81(4):471–482.
32. Braido F, Santus P, Corsico AG, et al. Chronic obstructive lung disease "expert system": validation of a predictive tool for assisting diagnosis. Int J Chron Obstruct Pulmon Dis. 2018;13:1747–1753. doi: 10.2147/COPD.S165533
33. Hafke-Dys H, Kuźnar-Kamińska B, Grzywalski T, et al. Artificial Intelligence Approach to the monitoring of respiratory sounds in asthmatic patients. Front Physiol. 2021;12:745635. doi: 10.3389/fphys.2021.745635
34. Buekers J, Theunis J, De Boever P, et al. Wearable finger pulse oximetry for continuous oxygen saturation measurements during daily home routines of patients with chronic obstructive pulmonary disease (COPD) over one week: observational study. JMIR mHealth uHealth. 2019 Jun 6;7(6):e12866. doi: 10.2196/12866
35. Mesquita R, Spina G, Pitta F, et al. Physical activity patterns and clusters in 1001 patients with COPD. Chron Respir Dis. 2017 Aug;14(3):256–269.
36. de Hond AAH, Leeuwenberg AM, Hooft L, et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. NPJ Digit Med. 2022 Jan 10;5(1):2. doi: 10.1038/s41746-021-00549-7
•• **Guidelines for the development of prediction models.**
37. Choi RY, Coyner AS, Kalpathy-Cramer J, et al. Introduction to machine learning, neural networks, and deep learning. Transl Vis Sci Technol. 2020;9(2):14. doi: 10.1167/tvst.9.2.14
38. Badillo S, Banfai B, Birzele F, et al. An introduction to machine learning. Clin Pharmacol Ther. 2020 Apr;107(4):871–885.
39. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. 2nd ed. Switzerland: Springer; 2019.
40. Naser MZ, Alavi AH. Error metrics and performance fitness indicators for artificial intelligence and machine learning in engineering and sciences. Archit Struct Constr. 2023 Nov 24;3(4):499–517.
41. Congalton RG. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens Environ. 1991 Jul 01;37(1):35–46.
42. Power M, Fell G, Wright M. Principles for high-quality, high-value testing. Evid Based Med. 2013 Feb;18(1):5–10. doi: 10.1136/eb-2012-100645
43. Hosmer DW, Lemeshow S. Applied Logistic Regression. 2nd ed. (New York): John Wiley and Sons; 2000.
44. MacNeil J, Loves RH, Aaron SD. Addressing the misdiagnosis of asthma in adults: where does it go wrong? Expert Rev Respir Med. 2016 Nov;10(11):1187–1198. doi: 10.1080/17476348.2016.1242415
45. Singh OP, Palaniappan R, Malarvili M. Automatic quantitative analysis of human respired carbon dioxide waveform for asthma and non-asthma classification using support vector machine. IEEE Access. 2018;6:55245–55256. doi: 10.1109/ACCESS.2018.2871091
46. Ullah R, Khan S, Ali H, et al. A comparative study of machine learning classifiers for risk prediction of asthma disease. Photodiagnosis Photodyn Ther. 2019 Dec;28:292–296.
47. Kisiel MA, Zhou X, Sundh J, et al. Data-driven questionnaire-based cluster analysis of asthma in Swedish adults. NPJ Prim Care Respir Med. 2020 Apr 6;30(1):14. doi: 10.1038/s41533-020-0168-0
48. Oh JH, Ahn KM, Chung SJ, et al. Usefulness of routine blood test-driven clusters for predicting acute exacerbation in patients with asthma. Respir Med. 2020 Aug;170:106042.
49. Zein JG, Wu CP, Attaway AH, et al. Novel Machine Learning Can Predict Acute Asthma Exacerbation. Chest. 2021 May;159(5):1747–1757.
50. Xiang Y, Ji H, Zhou Y, et al. Asthma exacerbation prediction and risk factor analysis based on a time-sensitive, attentive neural network: retrospective cohort study. J Med Internet Res. 2020 Jul 31;22(7):e16981. doi: 10.2196/16981
51. Luo G, He S, Stone BL, et al. Developing a model to predict hospital encounters for asthma in asthmatic patients: secondary analysis. JMIR Med Inform. 2020 Jan 21;8(1):e16080. doi: 10.2196/16080
52. Khasha R, Sepehri MM, Mahdaviani SA. An ensemble learning method for asthma control level detection with leveraging medical knowledge-based classifier and supervised learning. J Med Syst. 2019 Apr 26;43(6):158. doi: 10.1007/s10916-019-1259-8
53. Halner A, Beer S, Pullinger R, et al. Predicting treatment outcomes following an exacerbation of airways disease. PLoS One. 2021;16(8):e0254425. doi: 10.1371/journal.pone.0254425
54. Inbar O, Inbar O, Reuveny R, et al. A machine learning approach to the interpretation of cardiopulmonary exercise tests: development and validation. Pulm Med. 2021;2021:5516248–5516248. doi: 10.1155/2021/5516248
55. González G, Ash SY, Vegas-Sánchez-Ferrero G, et al. Disease staging and prognosis in smokers using deep learning in chest computed tomography. Am J Respir Crit Care Med. 2018 Jan 15;197(2):193–203. doi: 10.1164/rccm.201705-0860OC
56. Tang LYW, Coxson HO, Lam S, et al. 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 May;2(5):e259–e267.
57. Haider NS, Singh BK, Periyasamy R, et al. Respiratory sound based classification of chronic obstructive pulmonary disease: a risk stratification approach in machine learning paradigm. J Med Syst. 2019 Jun 28;43(8):255. doi: 10.1007/s10916-019-1388-0
58. Van Berkel JJ, Dallinga JW, Möller GM, et al. A profile of volatile organic compounds in breath discriminates COPD patients from controls. Respir Med. 2010 Apr;104(4):557–563.
59. B VA, Subramoniam M, Mathew L. Detection of COPD and lung cancer with electronic nose using ensemble learning methods. Clin Chim Acta. 2021 Dec 01;523:231–238.
60. Mariani S, Metting E, Lahr MMH, et al. Developing an ML pipeline for asthma and COPD: the case of a Dutch primary care service [10.1002/int.22568]. Int J Intell Syst. 2021;36(11):6763–6790. doi: 10.1002/int.22568
61. Amaral JL, Lopes AJ, Faria AC, et al. Machine learning algorithms and forced oscillation measurements to categorise the airway obstruction severity in chronic obstructive pulmonary disease. Comput Methods Programs Biomed. 2015 Feb;118(2):186–197.
62. Garcia-Aymerich J, Gómez FP, Benet M, et al. Identification and prospective validation of clinically relevant chronic obstructive pulmonary disease (COPD) subtypes. Thorax. 2011 May;66(5):430–7.

63. Marques A, Souto-Miranda S, Machado A, et al. COPD profiles and treatable traits using minimal resources: identification, decision tree and stability over time. Respir Res. 2022 Feb 14;23(1):30. doi: 10.1186/s12931-022-01954-6
64. Burgel P-R, Paillasseur J-L, Janssens W, et al. A simple algorithm for the identification of clinical COPD phenotypes. Eur Respir J. 2017;50(5):1701034. doi: 10.1183/13993003.01034-2017
65. Augustin IML, Spruit MA, Houben-Wilke S, et al. The respiratory physiome: clustering based on a comprehensive lung function assessment in patients with COPD. PLoS One. 2018;13(9): e0201593. doi: 10.1371/journal.pone.0201593
66. Cavailles A, Melloni B, Motola S, et al. Identification of patient profiles with high risk of hospital re-admissions for acute COPD exacerbations (AECOPD) in France using a machine learning model. Int J Chron Obstruct Pulmon Dis. 2020;15:949–962. 10.2147/COPD.S236787
67. Fernandez-Granero MA, Sanchez-Morillo D, Leon-Jimenez A. Computerised analysis of telemonitored respiratory sounds for predicting acute exacerbations of COPD. Sensors. 2015;15 (10):26978–26996. doi: 10.3390/s151026978
68. Aguilaniu B, Hess D, Kelkel E, et al. A machine learning approach to predict extreme inactivity in COPD patients using non-activity-related clinical data. PloS One. 2021;16(8):e0255977. doi: 10.1371/journal.pone.0255977
69. Moll M, Qiao D, Regan EA, et al. Machine learning and prediction of all-cause mortality in COPD. Chest. 2020 Sep;158 (3):952–964.
70. Luo L, Li J, Lian S, et al. Using machine learning approaches to predict high-cost chronic obstructive pulmonary disease patients in China. Health Inform J. 2020 Sep;26(3):1577–1598.
71. Ramspek CL, Jager KJ, Dekker FW, et al. External validation of prognostic models: what, why, how, when and where? Clin Kidney J. 2021;14(1):49–58. doi: 10.1093/ckj/sfaa188
72. Aung YYM, Wong DCS, Ting DSW. The promise of artificial intelligence: a review of the opportunities and challenges of artificial intelligence in healthcare. Br Med Bull. 2021 Sep 10;139(1):4–15. doi: 10.1093/bmb/ldab016
73. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit Health. 2021 Nov;3(11):e745–e750. doi: 10.1016/S2589-7500(21)00208-9
74. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell. 2019 May 01;1(5):206–215.
75. Martínez-Plumed F, Contreras-Ochando L, Ferri C, et al. CRISP-DM Twenty Years Later: from data mining processes to data science trajectories. IEEE Trans Knowl Data Eng. 2021;33(8):3048–3061. doi: 10.1109/TKDE.2019.2962680
76. Chapman P, Clinton J, Kerber R, et al. CRISP-DM 1.0: step-by-step data mining guide. USA: SPSS Inc; 2000.
77. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. 2015 Jan 7;350(4): g7594. doi: 10.1136/bmj.g7594