EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Review

Artificial Intelligence techniques in Asthma:A systematic review and critical appraisal of the existing literature

Konstantinos P. Exarchos, Maria Beltsiou, Chainti-Antonella Votti, Konstantinos Kostikas

Please cite this article as: Exarchos KP, Beltsiou M, Votti C-A, *et al*. Artificial Intelligence techniques in Asthma:A systematic review and critical appraisal of the existing literature. *Eur Respir J* 2020; in press (https://doi.org/10.1183/13993003.00521-2020).

Artificial Intelligence techniques in Asthma

A systematic review and critical appraisal of the existing literature

Konstantinos P. Exarchos, Maria Beltsiou, Chainti-Antonella Votti, Konstantinos Kostikas

Respiratory Medicine Department, School of Medicine, University of Ioannina, Ioannina, Greece

Abstract

Background: Artificial Intelligence (AI) when coupled with large amounts of well characterized data can yield models that are expected to facilitate clinical practice and contribute to the delivery of better care, especially in chronic diseases such as asthma.
Objective: The purpose of this paper is to review the utilization of AI techniques in all aspects of asthma research, i.e. from asthma screening and diagnosis, to patient classification and the overall asthma management and treatment, in order to identify trends, draw conclusions and discover potential gaps in the literature.
Methods: We conducted a systematic review of the literature using PubMed and DBLP from 1988 up to 2019, yielding 425 articles; after removing duplicate and irrelevant articles, 98 were further selected for detailed review.
Results: The resulting articles were organized in four categories, and subsequently compared based on a set of qualitative and quantitative factors. Overall, we observed an increasing adoption of AI techniques for asthma research, especially within the last decade.
Conclusions: AI is a scientific field that is in the spotlight, especially the last decade. In asthma there are already numerous studies, however, there are certain unmet needs that need to be further elucidated.

Introduction

Asthma is a common disease affecting an estimated 300 million individuals worldwide; in Europe, about 30 million children and adults less than 45 years old have asthma [1]. It is a major global health problem that imposes a substantial burden on patients, their families and the community. Asthma poses certain challenges that remain largely unmet despite the effort and the research in the respective fields, specifically: (i) There is no unanimous and widely applicable

diagnostic test for asthma, leading to significant underdiagnosis and overdiagnosis [2]. (ii) The **pathogenesis of asthma** is based on the process of gene-environment interaction, yet its specifics remain elusive; this field is currently in the spotlight in view of the new biologic treatments for asthma. (iii) **Asthma phenotypes** remain a controversial subject, due to the discordance in symptomatology, spirometry and response to treatment of individual patients. (iv) **Asthma exacerbations** play a crucial role in the course and management of the disease, incurring significant increase in direct and indirect costs [3].

As in other parts of medicine, there is an increasing interest in Artificial Intelligence (AI) methodologies to elucidate the aforementioned unmet needs of asthma. AI refers to the software that is able to make a machine intelligent such that it performs human tasks, i.e. process, learn and respond to information gained from data. The term is often used in combination with the term "Machine Learning (ML)" that refers to the process followed in order to make a machine learn how to perform a specific task, and in a similar manner as a human to perform better as the experience increases. Both AI and ML are data driven processes whereby the computer or the algorithm is presented with input data and the desired output and "learns" the inherent relations that lead from the input to the output. Similarly with AI and ML, Data Mining (DM) involves the computational and programming steps in order to "mine" large amounts of complex data for meaningful patterns and consequently knowledge. **Figure 1** depicts the steps of the DM process. There are two basic phases within the DM process: the training and the predicting phase. During the ***training phase*****,** the ML algorithm is fed with input data based on which a model is trained that captures the relations and patterns within the data. During the training phase the raw input data are subject to a series of preprocessing steps aiming to increase the quality of the data, identify the set of more informative features and omit potentially redundant or irrelevant information. Inherent to the training phase is the process of model evaluation where the parameters of the trained model are further fine-tuned in order to procure a well-trained model. In the ***predicting phase*** new instances of unknown data are fed as input to the previously trained model and the respective labels are predicted.

AI/ML in Medicine

Even though AI and ML exist as Computer Science domains for several years, they are two terms that have become radically and widely popular the last few years in a broad and non-specialized audience. This can be attributed to several reasons: their utilization in daily digital tasks especially pertaining to smartphones (e.g. mobile assistant, fingerprint scanner, personalized playlists, etc), the availability of AI and ML models to a wider audience with more user-friendly software, the need to discover new knowledge and analyze more effectively large and complex sources of data originating from various domains.

As described previously AI and ML models are largely dependent on the available data, and the healthcare domain is producing vast amounts of data that need to be mined for underlying knowledge. Such large and complex datasets incorporating various sources of data, e.g. clinical, imaging, genomic, proteomic, etc can be effectively analyzed with the available AI/ML techniques. In the supplement we provide a brief primer on AI/ML techniques in order to further facilitate reading of the manuscript. Imaging modalities such as CT and MRI scans that are used in clinical

practice can be effectively analyzed by ML algorithms [4, 5][6][7]. Genomic data is another source of enormous and complex information that is being used increasingly in the healthcare domain. Most of this data, e.g. SNPs, gene expression, etc produce large amounts of data that are impossible to comprehend; yet the systematic analysis of such data with ML techniques has brought about clinically meaningful knowledge for the benefit of patients [8][9]. The relatively recent boom of high quality wearable sensors is also producing huge amounts of time-series data that need to be mined efficiently in order to provide clinically relevant information [10, 11].

The distribution of data types analyzed with AI algorithms in the literature has been explored in a recent review article [12] suggesting that diagnostic imaging is the most widely employed data source in healthcare oriented applications of AI, while genomic data and electrodiagnosis constitute emerging data types that are equally appealing for analysis with AI. The authors further explored the leading diseases where AI algorithms have been employed in the literature, with cancer research being the top field where AI applications have been developed, followed by diseases of the nervous system as well as cardiovascular diseases [13]. In this analysis, respiratory diseases are way below with only mediocre adoption of AI techniques.

In the present manuscript we have systematically searched the literature for articles that employ AI or ML techniques in asthma, in an attempt to map the existing literature and identify gaps and areas of interest for future research. First, in the section “Literature review” we describe the methodological steps in order to acquire all relevant literature. Next, in the section “Machine Learning and asthma” we present our findings from the literature review, and the articles are organized in four major categories to facilitate the critical appraisal of the existing evidence.

Literature review

We systematically searched the literature until 18 May 2019 for articles using AI or ML techniques in asthma research. First, we searched DBLP, which is a computer science bibliography website using the term “asthma”. We maintained only journal articles posing no restriction regarding the year of publication. Next, we searched PubMed using the following terms: [“Artificial intelligence” AND asthma], [“Machine learning” AND asthma], [“Data mining” AND asthma], [“Decision Trees” AND asthma], [“Neural Network” AND asthma], [“Random Forests” AND asthma], [“Support Vector Machine” AND asthma]. The articles from both repositories were then merged and duplicates were removed. All articles were subsequently examined by the authors in order to exclude irrelevant ones; we also omitted articles not written in English. The aforementioned steps are shown in **Figure 2**, resulting eventually in 98 articles. Each of the 98 articles was then assigned to at least one out of the following four categories based on its content and purpose: (1) Asthma screening and diagnosis, (2) Patient classification, (3) Asthma management and monitoring, and (4) Asthma treatment.

In **Figure 3** we present the distribution of studies using AI/ML techniques for asthma research, over the course of approximately 30 years, from 1988 up to 2018. As expected during the first and second decade there is minimal use of such techniques for asthma, while from 2010 we observe a considerable and progressive increase.

Artificial Intelligence and asthma

In the sections that follow we present the articles retrieved in an organized manner, divided into four categories based on their content and purpose. Specifically, we have split the articles into the following four major contextual categories:

1. *Asthma screening and diagnosis*
2. *Patient classification*
3. *Asthma management and monitoring*
4. *Asthma treatment*

The articles from each category are summarized in a separate table where the respective studies can be compared by a set of qualitative and quantitative criteria or characteristics. These tables (**Table S3**, **Table S4** and **Table S5**) are available in the online supplement of the article. In the sections that follow we provide an overview of the articles comprising each category. Moreover, we have selected some of the most important studies from each category and provide more information. Our aim is to capture the most representative works, focusing on the ones that have been published within the last 5 years, as this is an emerging and developing field with rapid evolution.

In order to facilitate reading of the following sections, we hereby mention some terms that are commonly used in AI/ML. Specifically, ANN (Artificial Neural Network), RF (Random Forest), DT (Decision Tree), SVM (Support Vector Machine), LR (Logistic Regression), BN (Bayes Network), NB (Naive Bayes), k-NN (k Nearest Neighbors), SOM (Self Organizing Maps) and HMM (Hidden Markov Model) constitute classification algorithms; Se (Sensitivity), Sp (Specificity), Acc (Accuracy), ROC (Receiver Operating Characteristic) and AUC (Area Under ROC curve) are performance metrics used for the assessment of AI/ML algorithms. CV (Cross Validation) or its subtype called LOOCV (Leave One Out Cross Validation) are techniques used for AI/ML model validation. These terms are summarized in **Table 1** below. In the supplementary material we provide an exhaustive list of the abbreviations used throughout the manuscript, as well as a primer on AI/ML techniques.

ANN	Artificial Neural Network	**SOM**	Self Organizing Map
RF	Random Forest	**HMM**	Hidden Markov Model
DT	Decision Tree	**Se**	Sensitivity
SVM	Support Vector Machine	**Sp**	Specificity
LR	Logistic Regression	**Acc**	Accuracy
BN	Bayesian Network	**ROC**	Receiver Operating Curve
NB	Naive Bayes	**AUC**	Area Under ROC Curve
k-NN	k Nearest Neighbors	**LOOCV**	Leave One Out Cross Validation

Table 1: List of the most commonly used abbreviations in the manuscript.

1. Asthma screening and diagnosis

This category is the most populated one and contains 48 articles aiming for the screening or diagnosis of asthma. These studies are summarized in **Table S3** of the supplementary material. We observe that in terms of ML algorithms the majority of the studies (20 studies) employ ANNs or variations of ANNs, especially the earlier ones. SVMs are used in 8 studies, DTs or RFs are utilized in 11 studies, LR is used in 3 studies and k-NN in 2 studies. The remaining studies employ other ML algorithms such as HMM, fuzzy logic or NB. Overall, we observe that a limited number of ML algorithms are employed in the studies contained in the category 'Asthma screening and diagnosis', i.e. ANNs, SVMs, RF and DT. It should be noted that these ML algorithms are described in the accompanying Supplement, as well as some information regarding the evaluation of the reported results. Based on column 'Sample size' most of the studies employ tens or hundreds of patients and there are only a few studies that have enrolled larger patient cohorts (only four studies have enrolled more that 1000 patients).

As expected for the purpose of asthma diagnosis and screening, primarily clinical data have been employed; specifically, this data contains information from the medical history, pattern of symptoms, pulmonary function tests, lung sounds from auscultation etc. Clinical data are employed in 37 studies, out of which 12 explore features pertaining to lung or breath sounds. Similarly, there are studies in this category that exploit questionnaires as well as other clinical and epidemiological features in order to screen certain populations for asthma or identify patients that have a high probability of asthma. Some of the most recently published works employ genetic data (9 studies) in search of predisposing genetic traits for asthma. As for the evaluation methods, 23 studies used variations of cross validation techniques, of which 7 used LOOCV, 9 studies performed a train-test split and 9 studies used an independent test set. We have selected a few representative studies published within the last 5 years from the category 'Asthma screening and diagnosis' that we present briefly hereafter.

Oletic and Bilas [18] used a wearable sensor that recorded signals of respiratory sounds which were subsequently transferred to a smartphone. After certain signal manipulations, an HMM was utilized for respiratory sound classification, aiming primarily to detect wheezing. The resulting model yielded Acc=94.91%, Se=89.34% and Sp=96.28%. This study shows an emerging trend of smartphone employment in computationally intensive tasks such as the induction of ML algorithms in asthma.

Amaral and colleagues [19] explored the contribution of Forced Oscillation Technique (FOT) for the detection of airway obstruction, focusing specifically on patients with asthma. FOT is an oscillations-based technique that captures respiratory mechanics that can assess bronchial hyperresponsiveness in adults and children, and has been shown to be as sensitive as spirometry in detecting impairments to lung function due to smoking or exposure to occupational hazards [62]. It should be noted that FOT is a non-invasive technique that also has the advantage over conventional lung function tests that it does not require the performance of respiratory manoeuvres [62]. However, FOT should be used cautiously and as a complement to spirometry, since its interpretation and reference values remain controversial [63]. In their work Amaral et al [19] employed a series of ML algorithms using the FOT parameters as input in order to detect

airway obstruction. The best performance was achieved by a k-NN classifier that reached AUC=0.91.

In a methodologically different approach Kaur et al [26] utilized a Natural Language Processing (NLP) approach in order to mine health records and identify asthma diagnosis. The resulting algorithm was validated in a cohort of 427 patients and predicted asthma status with Se=86%, Sp=98%, PPV=88%, NPV=98%. Several approaches exist in the literature aiming to screen for asthma based on either the health record or the patient's prescriptions.

Singh et al [14] measure carbon dioxide waveforms from capnography in order to discriminate asthmatic and non-asthmatic patients. They extracted a series of features from the capnography signals from 30 non-asthmatic and 43 asthmatic patients; after applying feature selection, the remaining features were fed to a SVM which performed very well for the discrimination of the two classes (Acc=94.52%, Se=97.67%, Sp=90%). Capnography refers to the non-invasive measurement of the partial pressure of carbon dioxide (CO_2) in exhaled breath expressed as the CO_2 concentration over time. Changes in the CO_2 waveform (capnogram) or the end tidal CO_2 have been employed for disease diagnosis [64], assessment of disease severity as well as treatment response [65]. Based on the aforementioned results, the authors suggest that capnography may be a promising technique for diagnosing asthma, either alone or coupled with other features. The small dataset used in this study does not allow for proper evaluation of the proposed modality and further analyses in larger datasets are mandatory.

Another interesting work was recently published by Spathis and Vlamos [30] who developed a Decision Support System for the diagnosis of asthma and COPD. They used as input a set of clinical characteristics (e.g. age, sex, sputum production, chest pain, smoking, etc) as well as spirometry in order to detect asthma and COPD; the best performing algorithm in both cases was a Random forest classifier that resulted in Precision of 97.7% and 80.3% for the diagnosis of COPD and asthma, respectively. It should be noted that especially for COPD the results are quite encouraging given the fact that the employed input features are readily available during a regular pulmonology visit, yet the small number of patients does not allow for firm conclusions.

For a similar purpose as the previous work, Topalovic et al [41] employed spirometry and features from the patients' clinical profile in order to classify patients in 10 different conditions or states (healthy, asthma, COPD, other obstructive, hyperventilation, interstitial lung disease, neuromuscular disorder, pulmonary vascular disorders, upper airway obstruction). Compared to the evaluation by pulmonologists that resulted in correct diagnosis in approximately 38% of the subjects, the proposed ML algorithm that utilized a Decision Tree classifier achieved Acc=68%. The proposed algorithm performed better in the identification of spirometric patterns (obstructive, restrictive, mixed or normal) and in the most common conditions such as COPD and asthma.

Pandey et al [47] acquired nasal brushing samples from 190 patients with asthma and healthy controls and extracted RNA; the expression of 90 genes was recorded and fed to a Logistic Regression classifier which achieved an impressive AUC=0.994. Studies employing genomic data have recently emerged in the study of asthma but are gradually being used more widely, and can contribute to the pathogenesis of asthma at the molecular level. In a similar manner, Fang et al [50] analyzed gene expression data and came down to 62 genes that could serve as asthma biomarkers. Nasal brushing samples or gene expression data can often be acquired in a minimally invasive manner, nevertheless RNA extraction remains a costly technique.

Finally, the metabolome is another source of biomarkers that has recently been employed in a multitude of fields in medical research. Sinha et al [51] explored the Exhaled Breath Condensate (EBC) from 89 asthmatic subjects and 20 healthy controls and built a Random Forest classifier in order to differentiate between the two groups. The resulting classifier yielded Se=80% and Sp=75%. Same as before, EBC may be another promising field in the search for non-invasive asthma biomarkers, however this method needs further standardization prior to wider clinical application [66].

2. Patient classification

This category contains 31 studies that aim to classify patients into subgroups based on a series of characteristics. These subgroups refer to asthma severity, asthma phenotypes/endotypes or other classifications of patients. **Table S4** of the supplementary material shows a qualitative and quantitative comparison of these studies. In this category 9 studies employed DTs or RFs, 7 studies used ANNs, 3 studies utilized SVMs, 4 studies used LR and the remaining ones employed other ML techniques such as k-NN, BN, NB, etc.

The sample size, as expected, varies significantly among the studies. In terms of input data, 26 studies employ clinical data as input, whereas genomic data either alone or in combination with clinical information are used in 6 studies, especially in the most recent publications. Variations of the cross validation technique are primarily used (17 studies) for evaluating the proposed classifications schemes, out of which 10 studies use 10-fold CV and 3 studies employ the LOOCV method; 5 studies performed evaluation with an independent test set and 3 studies used the training-testing method.

It is noteworthy that this category ‘Patient classification’ is not the most populated one; however, it is the category that has significant overlap with the other categories. As noted before, every study based on its content could belong in more than one of the available categories. Studies in the category ‘Patient classification’ often belong to other categories as well. Specifically, the task of classifying patients into certain groups is often an important step in the studies even if there are other aims in the respective study.

Identifying subcategories within the broad category of ‘Patient classification’ is not easy. Roughly the studies in this category could be assigned into the following subcategories: i) asthma severity and ii) asthma phenotypes. Studies in the former subcategory feature a variety of inputs such as breath/respiratory sounds, asthma control and hospitalization frequency. Other studies explore the exacerbation severity and classify the patients according to the course of exacerbations or a set of clinical outcomes. Below we present a couple of the most recent and representative studies from the ‘asthma severity’ subcategory.

Van Vilet et al [68] explored the relationship between asthma control and exhaled biomarkers in a pediatric population. Specifically the authors explored the discriminatory ability of fractional nitric oxide (FeNO), Volatile Organic Compounds (VOCs) and cytokines/chemokines towards identifying children with persistently controlled and uncontrolled asthma. A cohort of 96 asthmatic children was followed up for a year and different features sets were fed as input to a Random Forest aiming to discriminate between the two patient groups. Using solely a set of VOCs

resulted in AUC=0.86, whereas the addition of the other two inputs did not lead to a more accurate classification.

Nabi and colleagues [69] analyzed wheeze sounds from 55 asthmatic patients in order to classify them into three severity classes i.e. mild, moderate and severe. An ensemble classifier yielded the highest PPV of 95%, pinpointing that tracheal related wheeze sounds were most sensitive and specific predictors of asthma severity levels.

Next, we focus on the second subcategory of the 'Patient classification' category, i.e. 'asthma phenotypes'. The studies in this subcategory either explore different patient classes based on a set of input features either genomic and/or clinical; therefore the patients are clustered based on their inherent characteristics. In the same subcategory there are studies that classify the employed patients based on their response to treatment. In the next few paragraphs we present some of the most important and recent studies from this subcategory.

Krautenbacher et al [72] combined a wide range of heterogeneous data, namely questionnaire, diagnostic, genotype, microarray, RT-qPCR, flow cytometry, and cytokine data in order to differentiate between three patient phenotypes. The phenotypes under consideration are healthy, mild-to-moderate allergic and nonallergic. The study focused on a pediatric population of 260 children. The most important variables for classifying childhood asthma phenotypes comprised novel identified genes, namely PKN2 (protein kinase N2), PTK2 (protein tyrosine kinase 2), and ALPP (alkaline phosphatase, placental). Similarly Fontanella et al [56] explored the relationship between allergic sensitization and asthma propensity; even though the study primarily aims to serve as a diagnostic tool for asthma, pairwise interactions between IgE components are used to predict clinical phenotypes.

Williams DeVane et al [80] utilized a completely data driven approach in order to identify asthma subtypes. The authors employed gene expression data, clinical covariates as well as certain disease indicators and devised a multi-step decision tree aiming to identify asthma endotypes aiming to facilitate the discovery of new mechanisms underlying asthma.

Wu et al [73] explored asthma phenotypes based on patients' response to corticosteroids, using an unsupervised multiview learning approach. The proposed work explored the contribution of 100 clinical, physiological, inflammatory, and demographic variables and was validated in a set of 346 adult asthmatic patients. The authors reported that patients with late-onset asthma and low lung function and high baseline eosinophilia showed the best corticosteroid responsiveness, whereas the poorest responsiveness was reported in young, obese females with severe airflow limitation and little eosinophilic inflammation. A similar approach is presented in the paper by Ross et al [88] where the authors proposed an ML algorithm in order to identify pediatric asthma phenotypes based on the patients' response to controller medication. Bronchodilator response and serum eosinophils were found to be the most predictive features of asthma control in the pediatric population under consideration.

3. Asthma management and monitoring

This category is also quite populated, featuring 40 studies that primarily deal with asthma exacerbations of asthma flare-ups. **Table S5** provides an overview of these studies. Regarding

ML algorithms, 12 studies employed DTs, RFs or variations of these algorithms; 11 studies utilized ANNs, 4 studies used SVMs, 3 studies employed BN/NB algorithms and 3 used LR.

Interestingly, in this category there are 11 studies employing more than 1000 records, where 5 of them analyze environmental data (e.g. air pollution). There are only 3 studies incorporating genomic data in this category, consequently the majority of the studies encompass either clinical data or environmental/meteorological data, or their combination (7 studies). CV was also the main method used for evaluation as reported in 21 studies, of which 2 used LOOCV; training-testing split was used in 8 studies and only 4 studies performed evaluation on an independent testing set. In this category we can identify two broad subcategories, namely asthma exacerbation prediction and asthma exacerbation management. The former category refers to models aiming to early identify an exacerbation while the latter contains models that predict the course of the exacerbation and the subsequent management.

Khasha et al [90] utilized experts' knowledge in an ensemble classifier in order to detect asthma control level yielding overall Acc=91.66%. The algorithm was developed with data collected from 96 asthmatic patients followed-up for a 9 month period. According to the authors, the aim of the proposed model is to serve as a real time preventive system for asthma control.

In a similar manner, Hosseini et al [92] proposed a platform for real-time assessment of asthma attack risk, based on a set of sensors capturing physiological and environmental data. The collected data are pipelined through a smartphone for analysis to an RF classifier which identified asthma attacks with an overall Acc=80.1%. In another work by Huffaker et al [101] nocturnal recordings of physiological data were obtained from a contactless bed sensor and fed to a RF model which yielded Acc=87.4%, Se=47.2% and Sp=96.3%, towards detecting asthma exacerbations. Similarly, for the prediction of asthma exacerbations Finkelstein et al [114] utilized telemonitoring data which were analyzed by an adaptive bayesian network resulting in perfect classification (i.e. Acc=100%, Se=100%, Sp=100%).

In a methodologically different approach Ram et al [95] mine a multitude of data coming from Google search interests, Twitter data and environmental data in order to early predict asthma-related emergency department visits; the resulting model yielded Precision=70%. Such systems could potentially serve as a means of public health surveillance in order to enhance proactiveness and efficiency of the Emergency Department. For the same purpose Khatri et al [102] developed an ANN model in order to predict peak demand days at the Emergency Department for chronic respiratory diseases.

Another important issue regarding an asthma exacerbation is the decision whether hospitalization is needed or not. Patel et al [107] proposed an algorithm based on gradient boosting machines that quantifies the overall risk and consequently the need for hospitalization is decided. The algorithm yielded AUC=0.84 and the following features were found to be more informative: vital signs, acuity, age, weight, socioeconomic status and weather-related features.

4. Asthma treatment

The last category contains studies utilizing machine learning algorithms for the overall asthma treatment. It is notable that this category contains only one article by Ross et al [88] which has also been mentioned in previous categories. The authors aimed to identify asthma

phenotypes based on their response to treatment and, thus, fine tune their patients' treatment. We have intentionally included this hardly populated category in order to highlight the gap in literature in terms of ML algorithms used for asthma treatment.

Discussion

Asthma research is gradually picking up on AI/ML techniques, following the overall trend of AI/ML adoption in healthcare related studies. Specifically, in **Figure 3** (Introduction section) we presented the distribution of studies using AI/ML techniques for asthma research, over the course of 30 years, i.e. from 1988 up to 2018. During the first and second decade there is minimal employment of such techniques for asthma, while from 2010 we observe a considerable and progressive increase. A similar trend has been observed regarding the utilization of AI/ML techniques in other healthcare domains, e.g. cancer research [13], whereas in the latter case the number of articles published in each year is almost ten times bigger.

In the 'Asthma screening and diagnosis' category we observe that the vast majority of studies have utilized relatively small numbers of patients. Only studies employing questionnaires contain richer patient sets. This observation poses an important question regarding the validity and robustness of the reported results.

As for the 'Patient classification' section, the studies employ relatively larger patient cohorts; nevertheless the reported evaluation metrics are encouraging but not quite perfect yet. Therefore, more data and further analyses are needed in order to obtain more definite answers.

The 'Asthma management and monitoring' category is quite heterogeneous in terms of the employed population sizes and the accuracy of the reported results. Specifically, we observe from the respective **Table S5** that the number of patients or records used in the studies vary significantly from just a couple up to thousands. This has to do with several factors: the type and cost of employed data (genomic, metabolomic, clinical, etc), the focus on specific populations and the scarcity of patients in each patient set, the quality and completeness of gathered information.

It is noteworthy that the last category ''Asthma treatment' contains one study, denoting the lack of research currently in this prospect with the employment of ML techniques. This can be attributed to the fact that treatment is primarily directed by published guidelines. However, it should be noted that in the field of asthma treatment there is considerable activity in the literature, especially with respect to biologics. According to our literature research there are currently no studies that exploit ML algorithms focused on the exploration of biologic treatments of asthma. Nevertheless, as the number of approved biologics increases, as also the number of eligible patients, such studies are expected to emerge. The profiles of super-responders to specific biologics currently remain largely elusive, and AI/ML could facilitate the discovery of such complex profiles. There is also an increasing interest in the reviewed literature towards severe asthma encompassing ML techniques, following the overall trend in asthma research.

Only a small fraction of the studies in the current review utilize large patient cohorts, and even fewer analyze complex data, where AI could be more useful; therefore, AI in asthma research still remains underused, or at least not exploited to its full potential. Furthermore, we observe that in terms of the quantitative and qualitative features we have compared the included studies, there are some similar patterns among them. Specifically, there is considerable utilization

of ANNs and DT, whereas in the most recent studies, RFs are being increasingly used. This trend is to be expected, since ANNs were widely employed in several medical fields due to their superior results. DTs are also quite common in health-related studies because they provide reasoning which is often regarded as cornerstone. It should be noted that there is a significant number of studies (i.e. 21) focusing on pediatric populations, whereas the rest include adults, denoting the burden based on age.

It should be highlighted that AI/ML techniques are particularly useful for the analysis of large complex datasets, encompassing heterogeneous sources of information. Asthma poses an ideal target for AI/ML utilization, as it is a chronic disease with patients being followed-up for several years and its perturbations can be detected from the cellular level, to the organ level and up to the organism level as a whole. Moreover, environmental factors play a key role in asthma pathogenesis and natural history, therefore, large scale environmental and meteorological data need to be analyzed in a complementary manner. Ideally, a theoretical asthma study should capture genomic, metabolomic, clinical and environmental data, in several consecutive time-slices from large and diverse patient cohorts, thus framing all potential asthma effects ranging in scale and time. The resulting highly complex and heterogeneous dataset should be mined with AI techniques aiming to gain new knowledge regarding asthma diagnosis, classification and treatment.

Conclusions

AI/ML is undeniably a scientific field that is in the spotlight, especially the last decade; its utilization in medical applications is on the rise, and subsequently there is growing interest in the respiratory field and asthma research, as denoted by the literature review conducted in the current work. Further progress is to be expected in respiratory research as more advanced ML techniques are gradually used, e.g. deep learning. Another issue that affects the combined research of asthma with AI/ML techniques is the fruitful communication between computer scientists and clinicians for the identification of the appropriate research questions. In order to deal with those questions more effectively large amounts of high quality and well characterized populations are needed. Finally, there is an unmet need in the identification of treatment responders to different therapeutic approaches, including the selection of an appropriate biologic treatment in severe asthma by predicting a patient's response based on phenotypic and endotypic characteristics. Artificial intelligence is here to stay in medicine, however there are certain open issues in asthma that need to be further elucidated.

References

1. Selroos O, Kupczyk M, Kuna P, Łacwik P, Bousquet J, Brennan D, Palkonen S, Contreras J, FitzGerald M, Hedlin G, Johnston SL, Louis R, Metcalf L, Walker S, Moreno-Galdó A, Papadopoulos NG, Rosado-Pinto J, Powell P, Haahtela T. National and regional asthma

programmes in Europe. *Eur. Respir. Rev.* 2015; 24: 474–483.

2. Aaron SD, Boulet LP, Reddel HK, Gershon AS. Underdiagnosis and Overdiagnosis of Asthma. *Am. J. Respir. Crit. Care Med.* 2018; 198: 1012–1020.

3. Bahadori K, Doyle-Waters MM, Marra C, Lynd L, Alasaly K, Swiston J, FitzGerald JM. Economic burden of asthma: a systematic review [Internet]. BMC Pulmonary Medicine 2009.Available from: http://dx.doi.org/10.1186/1471-2466-9-24.

4. Feng Y, Teh HS, Cai Y. Deep Learning for Chest Radiology: A Review [Internet]. Current Radiology Reports 2019.Available from: http://dx.doi.org/10.1007/s40134-019-0333-9.

5. McBee MP, Awan OA, Colucci AT, Ghobadi CW, Kadom N, Kansagra AP, Tridandapani S, Auffermann WF. Deep Learning in Radiology [Internet]. Academic Radiology 2018. p. 1472–1480Available from: http://dx.doi.org/10.1016/j.acra.2018.02.018.

6. Angelini E, Dahan S, Shah A. Unravelling Machine Learning – Insights in Respiratory Medicine [Internet]. European Respiratory Journal 2019. p. 1901216Available from: http://dx.doi.org/10.1183/13993003.01216-2019.

7. Young AL, Bragman FJS, Rangelov B, Han M, Galbán CJ, Lynch DA, Hawkes DJ, Alexander DC, Hurst JR, COPDGene Investigators. Disease Progression Modelling in Chronic Obstructive Pulmonary Disease (COPD). *Am. J. Respir. Crit. Care Med.* [Internet] 2019; Available from: http://dx.doi.org/10.1164/rccm.201908-1600OC.

8. Ho DSW, Schierding W, Wake M, Saffery R, O'Sullivan J. Machine Learning SNP Based Prediction for Precision Medicine. *Front. Genet.* 2019; 10: 267.

9. van IJzendoorn DGP, Szuhai K, Briaire-de Bruijn IH, Kostine M, Kuijjer ML, Bovée JVMG. Machine learning analysis of gene expression data reveals novel diagnostic and prognostic biomarkers and identifies therapeutic targets for soft tissue sarcomas. *PLoS Comput. Biol.* 2019; 15: e1006826.

10. Mauldin TR, Canby ME, Metsis V, Ngu AHH, Rivera CC. SmartFall: A Smartwatch-Based Fall Detection System Using Deep Learning. *Sensors* [Internet] 2018; 18Available from: http://dx.doi.org/10.3390/s18103363.

11. Tison GH, Sanchez JM, Ballinger B, Singh A, Olgin JE, Pletcher MJ, Vittinghoff E, Lee ES, Fan SM, Gladstone RA, Mikell C, Sohoni N, Hsieh J, Marcus GM. Passive Detection of Atrial Fibrillation Using a Commercially Available Smartwatch. *JAMA Cardiol* 2018; 3: 409–416.

12. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, Wang Y, Dong Q, Shen H, Wang Y. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol* 2017; 2: 230–243.

13. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput. Struct. Biotechnol. J.* 2015; 13: 8–17.

14. 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 [Internet]. IEEE Access 2018. p. 55245–55256Available from:

http://dx.doi.org/10.1109/access.2018.2871091.

15. Islam MA, Bandyopadhyaya I, Bhattacharyya P, Saha G. Multichannel lung sound analysis for asthma detection. *Comput. Methods Programs Biomed.* 2018; 159: 111–123.

16. Safdari R, Rezaei-Hachesu P, GhaziSaeedi M, Samad-Soltani T, Zolnoori M. Evaluation of Classification Algorithms vs Knowledge-Based Methods for Differential Diagnosis of Asthma in Iranian Patients [Internet]. International Journal of Information Systems in the Service Sector 2018. p. 22–35Available from: http://dx.doi.org/10.4018/ijisss.2018040102.

17. Gurbeta L, Badnjevic A, Maksimovic M, Omanovic-Miklicanin E, Sejdic E. A telehealth system for automated diagnosis of asthma and chronical obstructive pulmonary disease. *J. Am. Med. Inform. Assoc.* 2018; 25: 1213–1217.

18. Oletic D, Bilas V. Asthmatic Wheeze Detection From Compressively Sensed Respiratory Sound Spectra. *IEEE J Biomed Health Inform* 2018; 22: 1406–1414.

19. Amaral JLM, Lopes AJ, Veiga J, Faria ACD, Melo PL. High-accuracy detection of airway obstruction in asthma using machine learning algorithms and forced oscillation measurements. *Comput. Methods Programs Biomed.* 2017; 144: 113–125.

20. Mazić I, Bonković M, Džaja B. Two-level coarse-to-fine classification algorithm for asthma wheezing recognition in children's respiratory sounds [Internet]. Biomedical Signal Processing and Control 2015. p. 105–118Available from: http://dx.doi.org/10.1016/j.bspc.2015.05.002.

21. Chatzimichail E, Paraskakis E, Rigas A. Predicting Asthma Outcome Using Partial Least Square Regression and Artificial Neural Networks [Internet]. Advances in Artificial Intelligence 2013. p. 1–7Available from: http://dx.doi.org/10.1155/2013/435321.

22. Zolnoori M, Fazel Zarandi MH, Moin M, Heidarnezhad H, Kazemnejad A. Computer-aided intelligent system for diagnosing pediatric asthma. *J. Med. Syst.* 2012; 36: 809–822.

23. Chakraborty C, Mitra T, Mukherjee A, Ray AK. CAIDSA: Computer-aided intelligent diagnostic system for bronchial asthma [Internet]. Expert Systems with Applications 2009. p. 4958–4966Available from: http://dx.doi.org/10.1016/j.eswa.2008.06.025.

24. Zeng QT, Goryachev S, Weiss S, Sordo M, Murphy SN, Lazarus R. Extracting principal diagnosis, co-morbidity and smoking status for asthma research: evaluation of a natural language processing system. *BMC Med. Inform. Decis. Mak.* 2006; 6: 30.

25. Hirsch S, Shapiro J, Frank P. Use of an artificial neural network in estimating prevalence and assessing underdiagnosis of asthma [Internet]. Neural Computing & Applications 1997. p. 124–128Available from: http://dx.doi.org/10.1007/bf01501176.

26. Kaur H, Sohn S, Wi C-I, Ryu E, Park MA, Bachman K, Kita H, Croghan I, Castro-Rodriguez JA, Voge GA, Liu H, Juhn YJ. Automated chart review utilizing natural language processing algorithm for asthma predictive index. *BMC Pulm. Med.* 2018; 18: 34.

27. Alizadeh B, Safdari R, Zolnoori M, Bashiri A. Developing an Intelligent System for Diagnosis of Asthma Based on Artificial Neural Network. *Acta Inform. Med.* 2015; 23: 220–223.

28. Rother A-K, Schwerk N, Brinkmann F, Klawonn F, Lechner W, Grigull L. Diagnostic Support for Selected Paediatric Pulmonary Diseases Using Answer-Pattern Recognition in Questionnaires Based on Combined Data Mining Applications--A Monocentric Observational Pilot Study. *PLoS One* 2015; 10: e0135180.

29. Lozano M, Fiz JA, Jané R. Automatic Differentiation of Normal and Continuous Adventitious Respiratory Sounds Using Ensemble Empirical Mode Decomposition and Instantaneous Frequency. *IEEE J Biomed Health Inform* 2016; 20: 486–497.

30. Spathis D, Vlamos P. Diagnosing asthma and chronic obstructive pulmonary disease with machine learning. *Health Informatics J.* 2019; 25: 811–827.

31. Badnjevic A, Cifrek M, Koruga D, Osmankovic D. Neuro-fuzzy classification of asthma and chronic obstructive pulmonary disease. *BMC Med. Inform. Decis. Mak.* 2015; 15 Suppl 3: S1.

32. Lin B-S, Wu H-D, Chen S-J. Automatic Wheezing Detection Based on Signal Processing of Spectrogram and Back-Propagation Neural Network. *J. Healthc. Eng.* 2015; 6: 649–672.

33. Hashemi A, Arabalibeik H, Agin K. Classification of wheeze sounds using cepstral analysis and neural networks. *Stud. Health Technol. Inform.* 2012; 173: 161–165.

34. Dottorini T, Sole G, Nunziangeli L, Baldracchini F, Senin N, Mazzoleni G, Proietti C, Balaci L, Crisanti A. Serum IgE Reactivity Profiling in an Asthma Affected Cohort [Internet]. PLoS ONE 2011. p. e22319Available from: http://dx.doi.org/10.1371/journal.pone.0022319.

35. Montuschi P, Santonico M, Mondino C, Pennazza G, Mantini G, Martinelli E, Capuano R, Ciabattoni G, Paolesse R, Di Natale C, Barnes PJ, D'Amico A. Diagnostic performance of an electronic nose, fractional exhaled nitric oxide, and lung function testing in asthma. *Chest* 2010; 137: 790–796.

36. Tomida S, Hanai T, Koma N, Suzuki Y, Kobayashi T, Honda H. Artificial neural network predictive model for allergic disease using single nucleotide polymorphisms data. *J. Biosci. Bioeng.* 2002; 93: 470–478.

37. Hirsch S, Frank TL, Hazell M, Frank PI. Screening for asthma by population ranking: a validation study. *Ann. Epidemiol.* 2005; 15: 64–70.

38. Oud M. Lung function interpolation by means of neural-network-supported analysis of respiration sounds. *Med. Eng. Phys.* 2003; 25: 309–316.

39. Hirsch S, Shapiro JL, Turega MA, Frank TL, Niven RM, Frank PI. Using a neural network to screen a population for asthma. *Ann. Epidemiol.* 2001; 11: 369–376.

40. Malmberg LP, Kallio K, Haltsonen S, Katila T, Sovijärvi AR. Classification of lung sounds in patients with asthma, emphysema, fibrosing alveolitis and healthy lungs by using self-organizing maps. *Clin. Physiol.* 1996; 16: 115–129.

41. Topalovic M, Laval S, Aerts J-M, Troosters T, Decramer M, Janssens W, Belgian Pulmonary Function Study investigators. Automated Interpretation of Pulmonary Function Tests in Adults with Respiratory Complaints. *Respiration* 2017; 93: 170–178.

42. Metting EI, In 't Veen JCCM, Dekhuijzen PNR, van Heijst E, Kocks JWH, Muilwijk-Kroes

JB, Chavannes NH, van der Molen T. Development of a diagnostic decision tree for obstructive pulmonary diseases based on real-life data. *ERJ Open Res* [Internet] 2016; 2Available from: http://dx.doi.org/10.1183/23120541.00077-2015.

43. Oletic D, Arsenali B, Bilas V. Low-power wearable respiratory sound sensing. *Sensors* 2014; 14: 6535–6566.

44. Prosperi MC, Marinho S, Simpson A, Custovic A, Buchan IE. Predicting phenotypes of asthma and eczema with machine learning. *BMC Med. Genomics* 2014; 7 Suppl 1: S7.

45. Prosperi MCF, Belgrave D, Buchan I, Simpson A, Custovic A. Challenges in interpreting allergen microarrays in relation to clinical symptoms: a machine learning approach. *Pediatr. Allergy Immunol.* 2014; 25: 71–79.

46. Bahoura M. Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes. *Comput. Biol. Med.* 2009; 39: 824–843.

47. Pandey G, Pandey OP, Rogers AJ, Ahsen ME, Hoffman GE, Raby BA, Weiss ST, Schadt EE, Bunyavanich S. A Nasal Brush-based Classifier of Asthma Identified by Machine Learning Analysis of Nasal RNA Sequence Data. *Sci. Rep.* 2018; 8: 8826.

48. Bokov P, Mahut B, Flaud P, Delclaux C. Wheezing recognition algorithm using recordings of respiratory sounds at the mouth in a pediatric population [Internet]. Computers in Biology and Medicine 2016. p. 40–50Available from: http://dx.doi.org/10.1016/j.compbiomed.2016.01.002.

49. Afzal Z, Engelkes M, Verhamme KMC, Janssens HM, Sturkenboom MCJM, Kors JA, Schuemie MJ. Automatic generation of case-detection algorithms to identify children with asthma from large electronic health record databases. *Pharmacoepidemiol. Drug Saf.* 2013; 22: 826–833.

50. Fang F, Pan J, Li Y, Li Y, Feng X, Wang J. Identification of potential transcriptomic markers in developing asthma: An integrative analysis of gene expression profiles. *Mol. Immunol.* 2017; 92: 38–44.

51. Sinha A, Desiraju K, Aggarwal K, Kutum R, Roy S, Lodha R, Kabra SK, Ghosh B, Sethi T, Agrawal A. Exhaled breath condensate metabolome clusters for endotype discovery in asthma. *J. Transl. Med.* 2017; 15: 262.

52. Panganiban RP, Wang Y, Howrylak J, Chinchilli VM, Craig TJ, August A, Ishmael FT. Circulating microRNAs as biomarkers in patients with allergic rhinitis and asthma. *J. Allergy Clin. Immunol.* 2016; 137: 1423–1432.

53. Burge PS, Pantin CF, Newton DT, Gannon PF, Bright P, Belcher J, McCoach J, Baldwin DR, Burge CB. Development of an expert system for the interpretation of serial peak expiratory flow measurements in the diagnosis of occupational asthma. Midlands Thoracic Society Research Group [Internet]. Occupational and Environmental Medicine 1999. p. 758–764Available from: http://dx.doi.org/10.1136/oem.56.11.758.

54. Oud M, Dooijes EH, van der Zee JS. Asthmatic airways obstruction assessment based on detailed analysis of respiratory sound spectra. *IEEE Trans. Biomed. Eng.* 2000; 47: 1450–1455.

55. Rietveld S, Oud M, Dooijes EH. Classification of asthmatic breath sounds: preliminary results of the classifying capacity of human examiners versus artificial neural networks. *Comput. Biomed. Res.* 1999; 32: 440–448.

56. Fontanella S, Frainay C, Murray CS, Simpson A, Custovic A. Machine learning to identify pairwise interactions between specific IgE antibodies and their association with asthma: A cross-sectional analysis within a population-based birth cohort. *PLoS Med.* 2018; 15: e1002691.

57. Rodrigo-Muñoz JM, Cañas JA, Sastre B, Rego N, Greif G, Rial M, Mínguez P, Mahíllo-Fernández I, Fernández-Nieto M, Mora I, Barranco P, Quirce S, Sastre J, Del Pozo V. Asthma diagnosis using integrated analysis of eosinophil microRNAs. *Allergy* 2019; 74: 507–517.

58. Farion KJ, Wilk S, Michalowski W, O'Sullivan D, Sayyad-Shirabad J. Comparing predictions made by a prediction model, clinical score, and physicians: pediatric asthma exacerbations in the emergency department. *Appl. Clin. Inform.* 2013; 4: 376–391.

59. Kazemi M, Bala Krishnan M, Aik Howe T. Frequency analysis of capnogram signals to differentiate asthmatic and non-asthmatic conditions using radial basis function neural networks. *Iran. J. Allergy Asthma Immunol.* 2013; 12: 236–246.

60. Grassi M, Villani S, Marinoni A, ECRHS Group. European Community Respiratory Health Survey. Classification methods for the identification of "case" in epidemiological diagnosis of asthma. *Eur. J. Epidemiol.* 2001; 17: 19–29.

61. Wang C-H, Liu B-J, Wu LS-H. The association forecasting of 13 variants within seven asthma susceptibility genes on 3 serum IgE groups in Taiwanese population by integrating of adaptive neuro-fuzzy inference system (ANFIS) and classification analysis methods. *J. Med. Syst.* 2012; 36: 175–185.

62. Oostveen E, MacLeod D, Lorino H, Farré R, Hantos Z, Desager K, Marchal F, ERS Task Force on Respiratory Impedance Measurements. The forced oscillation technique in clinical practice: methodology, recommendations and future developments. *Eur. Respir. J.* 2003; 22: 1026–1041.

63. Shirai T, Kurosawa H. Clinical Application of the Forced Oscillation Technique [Internet]. Internal Medicine 2016. p. 559–566Available from: http://dx.doi.org/10.2169/internalmedicine.55.5876.

64. Chebl RB, Madden B, Belsky J, Harmouche E, Yessayan L. Diagnostic value of end tidal capnography in patients with hyperglycemia in the emergency department [Internet]. BMC Emergency Medicine 2016.Available from: http://dx.doi.org/10.1186/s12873-016-0072-7.

65. Paiva EF, Paxton JH, O'Neil BJ. The use of end-tidal carbon dioxide (ETCO) measurement to guide management of cardiac arrest: A systematic review. *Resuscitation* 2018; 123: 1–7.

66. Horváth I, Barnes PJ, Loukides S, Sterk PJ, Högman M, Olin A-C, Amann A, Antus B, Baraldi E, Bikov A, Boots AW, Bos LD, Brinkman P, Bucca C, Carpagnano GE, Corradi M, Cristescu S, de Jongste JC, Dinh-Xuan A-T, Dompeling E, Fens N, Fowler S, Hohlfeld JM, Holz O, Jöbsis Q, Van De Kant K, Knobel HH, Kostikas K, Lehtimäki L, Lundberg J, et al. A European Respiratory Society technical standard: exhaled biomarkers in lung disease. *Eur.*

Respir. J. [Internet] 2017; 49Available from: http://dx.doi.org/10.1183/13993003.00965-2016.

67. Tomita Y, Tomida S, Hasegawa Y, Suzuki Y, Shirakawa T, Kobayashi T, Honda H. Artificial neural network approach for selection of susceptible single nucleotide polymorphisms and construction of prediction model on childhood allergic asthma. *BMC Bioinformatics* 2004; 5: 120.

68. Van Vliet D, Smolinska A, Jöbsis Q, Rosias PPR, Muris JWM, Dallinga JW, van Schooten FJ, Dompeling E. Association between exhaled inflammatory markers and asthma control in children. *J. Breath Res.* 2016; 10: 016014.

69. Nabi FG, Sundaraj K, Lam CK, Palaniappan R. Characterization and classification of asthmatic wheeze sounds according to severity level using spectral integrated features. *Comput. Biol. Med.* 2019; 104: 52–61.

70. Zolnoori M, Zarandi MHF, Moin M, Teimorian S. Fuzzy rule-based expert system for assessment severity of asthma. *J. Med. Syst.* 2012; 36: 1707–1717.

71. Farion K, Michalowski W, Wilk S, O'Sullivan D, Matwin S. A Tree-Based Decision Model to Support Prediction of the Severity of Asthma Exacerbations in Children [Internet]. Journal of Medical Systems 2010. p. 551–562Available from: http://dx.doi.org/10.1007/s10916-009-9268-7.

72. Krautenbacher N, Flach N, Böck A, Laubhahn K, Laimighofer M, Theis FJ, Ankerst DP, Fuchs C, Schaub B. A strategy for high-dimensional multivariable analysis classifies childhood asthma phenotypes from genetic, immunological, and environmental factors. *Allergy* 2019; 74: 1364–1373.

73. Wu W, Bang S, Bleecker ER, Castro M, Denlinger L, Erzurum SC, Fahy JV, Fitzpatrick AM, Gaston BM, Hastie AT, Israel E, Jarjour NN, Levy BD, Mauger DT, Meyers DA, Moore WC, Peters M, Phillips BR, Phipatanakul W, Sorkness RL, Wenzel SE. Multiview Cluster Analysis Identifies Variable Corticosteroid Response Phenotypes in Severe Asthma. *Am. J. Respir. Crit. Care Med.* 2019; 199: 1358–1367.

74. Kuo C-HS, Pavlidis S, Loza M, Baribaud F, Rowe A, Pandis I, Hoda U, Rossios C, Sousa A, Wilson SJ, Howarth P, Dahlen B, Dahlen S-E, Chanez P, Shaw D, Krug N, Sandström T, De Meulder B, Lefaudeux D, Fowler S, Fleming L, Corfield J, Auffray C, Sterk PJ, Djukanovic R, Guo Y, Adcock IM, Chung KF, U-BIOPRED Project Team ‡. A Transcriptome-driven Analysis of Epithelial Brushings and Bronchial Biopsies to Define Asthma Phenotypes in U-BIOPRED. *Am. J. Respir. Crit. Care Med.* 2017; 195: 443–455.

75. Wu J, Prosperi MCF, Simpson A, Hollams EM, Sly PD, Custovic A, Holt PG. Relationship between cytokine expression patterns and clinical outcomes: two population-based birth cohorts. *Clin. Exp. Allergy* 2015; 45: 1801–1811.

76. Wu W, Bleecker E, Moore W, Busse WW, Castro M, Chung KF, Calhoun WJ, Erzurum S, Gaston B, Israel E, Curran-Everett D, Wenzel SE. Unsupervised phenotyping of Severe Asthma Research Program participants using expanded lung data. *J. Allergy Clin. Immunol.* 2014; 133: 1280–1288.

77. Lazic N, Roberts G, Custovic A, Belgrave D, Bishop CM, Winn J, Curtin JA, Hasan Arshad

S, Simpson A. Multiple atopy phenotypes and their associations with asthma: similar findings from two birth cohorts. *Allergy* 2013; 68: 764–770.

78. Brasier AR, Victor S, Ju H, Busse WW, Curran-Everett D, Bleecker E, Castro M, Chung KF, Gaston B, Israel E, Wenzel SE, Erzurum SC, Jarjour NN, Calhoun WJ. Predicting intermediate phenotypes in asthma using bronchoalveolar lavage-derived cytokines. *Clin. Transl. Sci.* 2010; 3: 147–157.

79. Bureau A, Dupuis J, Falls K, Lunetta KL, Hayward B, Keith TP, Van Eerdewegh P. Identifying SNPs predictive of phenotype using random forests. *Genet. Epidemiol.* 2005; 28: 171–182.

80. Williams-DeVane CR, Reif DM, Hubal EC, Bushel PR, Hudgens EE, Gallagher JE, Edwards SW. Decision tree-based method for integrating gene expression, demographic, and clinical data to determine disease endotypes. *BMC Syst. Biol.* 2013; 7: 119.

81. Tsai C-L, Clark S, Camargo CA Jr. Risk stratification for hospitalization in acute asthma: the CHOP classification tree. *Am. J. Emerg. Med.* 2010; 28: 803–808.

82. Zolnoori M, Fazel Zarandi MH, Moin M, Taherian M. Fuzzy rule-based expert system for evaluating level of asthma control. *J. Med. Syst.* 2012; 36: 2947–2958.

83. Messinger AI, Bui N, Wagner BD, Szefler SJ, Vu T, Deterding RR. Novel pediatric-automated respiratory score using physiologic data and machine learning in asthma. *Pediatr. Pulmonol.* 2019; 54: 1149–1155.

84. Işik AH, Güler I, Sener MU. A low-cost mobile adaptive tracking system for chronic pulmonary patients in home environment. *Telemed. J. E. Health.* 2013; 19: 24–30.

85. Park H-W, Song W-J, Kim S-H, Park H-K, Kim S-H, Kwon YE, Kwon H-S, Kim T-B, Chang Y-S, Cho Y-S, Lee B-J, Jee Y-K, Jang A-S, Nahm D-H, Park J-W, Yoon HJ, Cho Y-J, Choi BW, Moon H-B, Cho S-H. Classification and implementation of asthma phenotypes in elderly patients. *Ann. Allergy Asthma Immunol.* 2015; 114: 18–22.

86. Himes BE, Kohane IS, Ramoni MF, Weiss ST. Characterization of patients who suffer asthma exacerbations using data extracted from electronic medical records. *AMIA Annu. Symp. Proc.* 2008; : 308–312.

87. Belgrave DCM, Granell R, Simpson A, Guiver J, Bishop C, Buchan I, Henderson AJ, Custovic A. Developmental profiles of eczema, wheeze, and rhinitis: two population-based birth cohort studies. *PLoS Med.* 2014; 11: e1001748.

88. Ross MK, Yoon J, van der Schaar A, van der Schaar M. Discovering Pediatric Asthma Phenotypes on the Basis of Response to Controller Medication Using Machine Learning. *Ann. Am. Thorac. Soc.* 2018; 15: 49–58.

89. Zolnoori M, Zarandi MHF, Moin M. Application of intelligent systems in asthma disease: designing a fuzzy rule-based system for evaluating level of asthma exacerbation. *J. Med. Syst.* 2012; 36: 2071–2083.

90. 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; 43: 158.

91. Gardeux V, Berghout J, Achour I, Schissler AG, Li Q, Kenost C, Li J, Shang Y, Bosco A, Saner D, Halonen MJ, Jackson DJ, Li H, Martinez FD, Lussier YA. A genome-by-environment interaction classifier for precision medicine: personal transcriptome response to rhinovirus identifies children prone to asthma exacerbations. *J. Am. Med. Inform. Assoc.* 2017; 24: 1116–1126.

92. Hosseini A, Buonocore CM, Hashemzadeh S, Hojaiji H, Kalantarian H, Sideris C, Bui AAT, King CE, Sarrafzadeh M. Feasibility of a Secure Wireless Sensing Smartwatch Application for the Self-Management of Pediatric Asthma. *Sensors* [Internet] 2017; 17Available from: http://dx.doi.org/10.3390/s17081780.

93. Toti G, Vilalta R, Lindner P, Lefer B, Macias C, Price D. Analysis of correlation between pediatric asthma exacerbation and exposure to pollutant mixtures with association rule mining. *Artif. Intell. Med.* 2016; 74: 44–52.

94. Luo G, Stone BL, Fassl B, Maloney CG, Gesteland PH, Yerram SR, Nkoy FL. Predicting asthma control deterioration in children. *BMC Med. Inform. Decis. Mak.* 2015; 15: 84.

95. Ram S, Zhang W, Williams M, Pengetnze Y. Predicting asthma-related emergency department visits using big data. *IEEE J Biomed Health Inform* 2015; 19: 1216–1223.

96. Chatzimichail E, Paraskakis E, Sitzimi M, Rigas A. An intelligent system approach for asthma prediction in symptomatic preschool children. *Comput. Math. Methods Med.* 2013; 2013: 240182.

97. Lee C-H, Chen JC-Y, Tseng VS. A novel data mining mechanism considering bio-signal and environmental data with applications on asthma monitoring. *Comput. Methods Programs Biomed.* 2011; 101: 44–61.

98. Kimes D, Nelson R, Levine E, Weiss S, Bollinger ME, Blaisdell C. Predicting paediatric asthma hospital admissions and ED visits [Internet]. Neural Computing & Applications 2003. p. 10–17Available from: http://dx.doi.org/10.1007/s00521-003-0366-z.

99. Al-Momani O, Gharaibeh KM. Effect of wireless channels on detection and classification of asthma attacks in wireless remote health monitoring systems. *Int. J. Telemed. Appl.* 2014; 2014: 816369.

100. Goto T, Camargo CA Jr, Faridi MK, Yun BJ, Hasegawa K. Machine learning approaches for predicting disposition of asthma and COPD exacerbations in the ED. *Am. J. Emerg. Med.* 2018; 36: 1650–1654.

101. Huffaker MF, Carchia M, Harris BU, Kethman WC, Murphy TE, Sakarovitch CCD, Qin F, Cornfield DN. Passive Nocturnal Physiologic Monitoring Enables Early Detection of Exacerbations in Children with Asthma. A Proof-of-Concept Study. *Am. J. Respir. Crit. Care Med.* 2018; 198: 320–328.

102. Khatri KL, Tamil LS. Early Detection of Peak Demand Days of Chronic Respiratory Diseases Emergency Department Visits Using Artificial Neural Networks. *IEEE J Biomed Health Inform* 2018; 22: 285–290.

103. Moustris KP, Douros K, Nastos PT, Larissi IK, Anthracopoulos MB, Paliatsos AG, Priftis KN. Seven-days-ahead forecasting of childhood asthma admissions using artificial neural networks in Athens, Greece. *Int. J. Environ. Health Res.* 2012; 22: 93–104.

104. Jaing JT, Sepulveda JA, Casillas AM. Novel computer-based assessment of asthma strategies in inner-city children. *Ann. Allergy Asthma Immunol.* 2001; 87: 230–237.

105. Nutman A, Solomon Y, Mendel S, Nutman J, Hines E, Topilsky M, Kivity S. The use of a neural network for studying the relationship between air pollution and asthma-related emergency room visits. *Respir. Med.* 1998; 92: 1199–1202.

106. Moseholm L, Taudorf E, Frøsig A. Pulmonary function changes in asthmatics associated with low-level SO2 and NO2 air pollution, weather, and medicine intake. An 8-month prospective study analyzed by neural networks. *Allergy* 1993; 48: 334–344.

107. Patel SJ, Chamberlain DB, Chamberlain JM. A Machine Learning Approach to Predicting Need for Hospitalization for Pediatric Asthma Exacerbation at the Time of Emergency Department Triage. *Acad. Emerg. Med.* 2018; 25: 1463–1470.

108. Kerem E, Tibshirani R, Canny G, Bentur L, Reisman J, Schuh S, Stein R, Levison H. Predicting the need for hospitalization in children with acute asthma. *Chest* 1990; 98: 1355–1361.

109. Finkelstein J, Wood J. Predicting asthma exacerbations using artificial intelligence. *Stud. Health Technol. Inform.* 2013; 190: 56–58.

110. Xu M, Tantisira KG, Wu A, Litonjua AA, Chu J-H, Himes BE, Damask A, Weiss ST. Genome Wide Association Study to predict severe asthma exacerbations in children using random forests classifiers. *BMC Med. Genet.* 2011; 12: 90.

111. Deng H, Urman R, Gilliland FD, Eckel SP. Understanding the importance of key risk factors in predicting chronic bronchitic symptoms using a machine learning approach. *BMC Med. Res. Methodol.* 2019; 19: 70.

112. Luo L, Liao C, Zhang F, Zhang W, Li C, Qiu Z, Huang D. Applicability of internet search index for asthma admission forecast using machine learning. *Int. J. Health Plann. Manage.* [Internet] 2018; Available from: http://dx.doi.org/10.1002/hpm.2525.

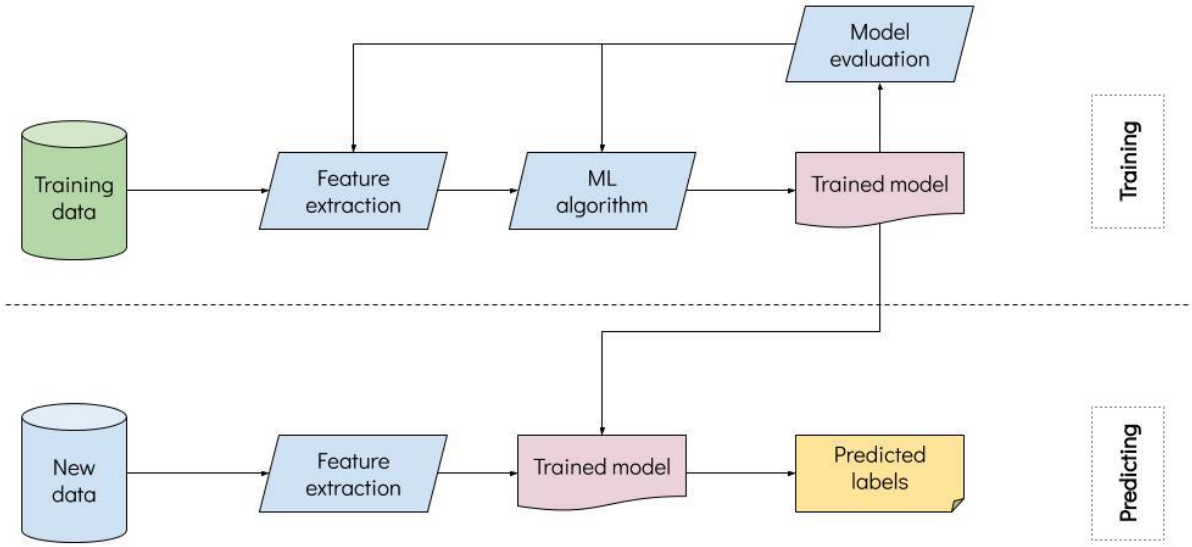
113. Das LT, Abramson EL, Stone AE, Kondrich JE, Kern LM, Grinspan ZM. Predicting frequent emergency department visits among children with asthma using EHR data. *Pediatr. Pulmonol.* 2017; 52: 880–890.

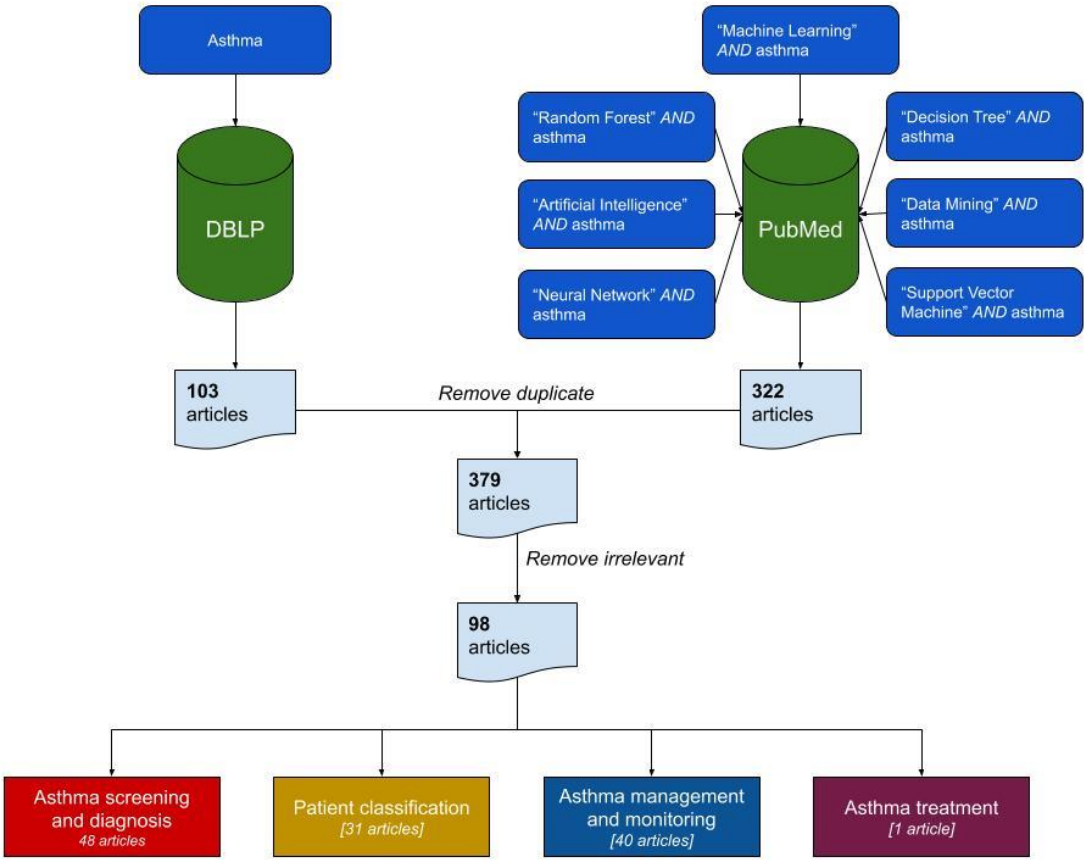
114. Finkelstein J, Jeong IC. Machine learning approaches to personalize early prediction of asthma exacerbations. *Ann. N. Y. Acad. Sci.* 2017; 1387: 153–165.

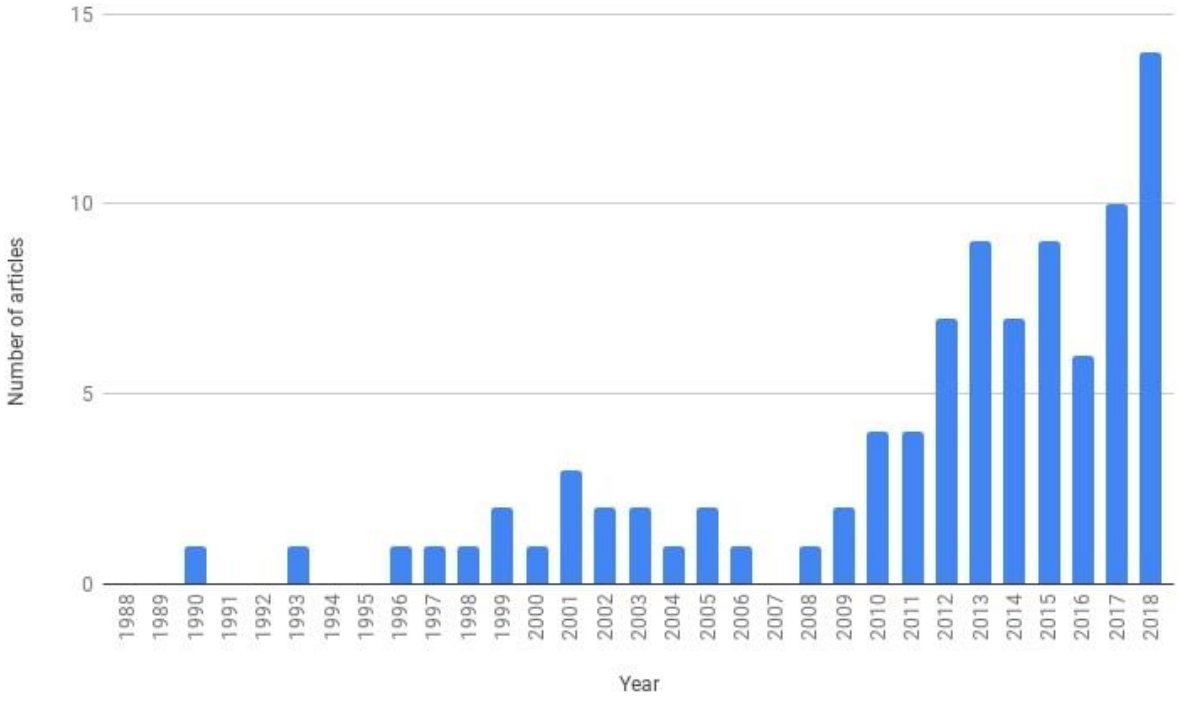
115. Sterling M, Rhee H, Bocko M. Automated Cough Assessment on a Mobile Platform. *J. Med. Eng. Technol.* [Internet] 2014; 2014Available from: http://dx.doi.org/10.1155/2014/951621.

116. Pifferi M, Bush A, Pioggia G, Di Cicco M, Chinellato I, Bodini A, Macchia P, Boner AL. Monitoring asthma control in children with allergies by soft computing of lung function and exhaled nitric oxide. *Chest* 2011; 139: 319–327.

Figure 1: Flowchart of the Data Mining (DM) process.
Figure 2: Flowchart of the literature search.
Figure 3: Distribution of articles published per year, that employ AI/ML techniques for asthma research.


Model evaluation
Training data
Feature extraction
ML algorithm
Trained model
Training
New data
Feature extraction
Trained model
Predicted labels
Predicting


Asthma
"Machine Learning" AND asthma
"Random Forest" AND asthma
"Decision Tree" AND asthma
"Artificial Intelligence" AND asthma
"Data Mining" AND asthma
"Neural Network" AND asthma
"Support Vector Machine" AND asthma
DBLP
PubMed
103 articles
Remove duplicate
322 articles
379 articles
Remove irrelevant
98 articles
Asthma screening and diagnosis
48 articles
Patient classification
[31 articles]
Asthma management and monitoring
[40 articles]
Asthma treatment
[1 article]


15
10
5
0
Number of articles
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
Year

Supplementary material for the article “Artificial Intelligence techniques in Asthma: A systematic review and critical appraisal of the existing literature”

Table of contents

Introduction to Artificial Intelligence

The aim of this supplement is to serve as an introduction in Artificial Intelligence. First, we provide a table of abbreviations where the reader may refer, in order to facilitate the comprehension of the manuscript. In the next section “AI/Ml flavors” we describe the Data Mining process and present a rough categorization of AI/ML techniques based on the learning process. Next, we briefly describe the most commonly employed classification algorithms, especially the ones that are frequently used in medicine oriented problems.

Table of Abbreviations

In this section (**Table S1**) we provide a list of the most commonly used abbreviations pertaining to Artificial Intelligence that are frequently used throughout the manuscript.

Table S1 contains a list of the abbreviations used throughout the main body of this manuscript.

AI	Artificial Intelligence
ML	Machine Learning
DM	Data mining
ANN	Artificial Neural Networks
RF	Random Forest
DT	Decision Tree
PCA	Principal Component Analysis
SVM	Support Vector Machines
LR	Logistic Regression
BN	Bayesian Network
HMM	Hidden Markov Model
k-NN	k Nearest Neighbors
SOM	Self Organizing Map
GMM	Gaussian Mixture Model
NB	Naive Bayes
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
Se	Sensitivity
TPR	True Positive Rate

Sp	Specificity
TNR	True Negative Rate
Acc	Accuracy
ROC	Receiver Operating Characteristic
AUC	Area Under ROC Curve
PPV	Positive Predictive Value
NPV	Negative Predictive Value
LOOCV	Leave One Out Cross Validation

Table S1: Table of the most frequently used abbreviations in this section.

AI/ML “flavors”

Artificial Intelligence (AI) refers to the software that is able to make a machine intelligent such that it performs human tasks, i.e. process, learn and respond to information gained from data; whereas Machine Learning (ML) is the process followed in order to make a machine learn how to perform a specific task, and in a similar manner as a human to perform better as the experience increases. Both AI and ML are data driven processes whereby the computer or the algorithm is presented with input data and the desired output and subsequently “learns” the inherent relations that lead from the input to the output. This is a completely different approach compared to a traditional computer programme where input data are fed and based on a set of extremely precise predefined instructions the computer returns a specific outcome. Similarly with AI and ML, Data Mining (DM) involves the computational and programming steps in order to “mine” large amounts of complex data for meaningful patterns and consequently knowledge. **Figure S1** depicts the steps of the DM process. There are roughly two basic phases within the DM process: i) during the **training phase,** the ML algorithm is fed with input data based on which a model is trained that captures the relations and inherent patterns within the data. During the training phase the raw input data are subject to a series of preprocessing steps aiming to increase the quality of the data, identify the set of more informative features and omit potentially redundant or irrelevant information. Inherent to the training phase is the process of model evaluation where the parameters of the trained model are further fine-tuned in order to procure a well-trained model. ii) In the **predicting phase** new instances of unknown data are fed as input to the previously trained model and the respective labels are predicted.

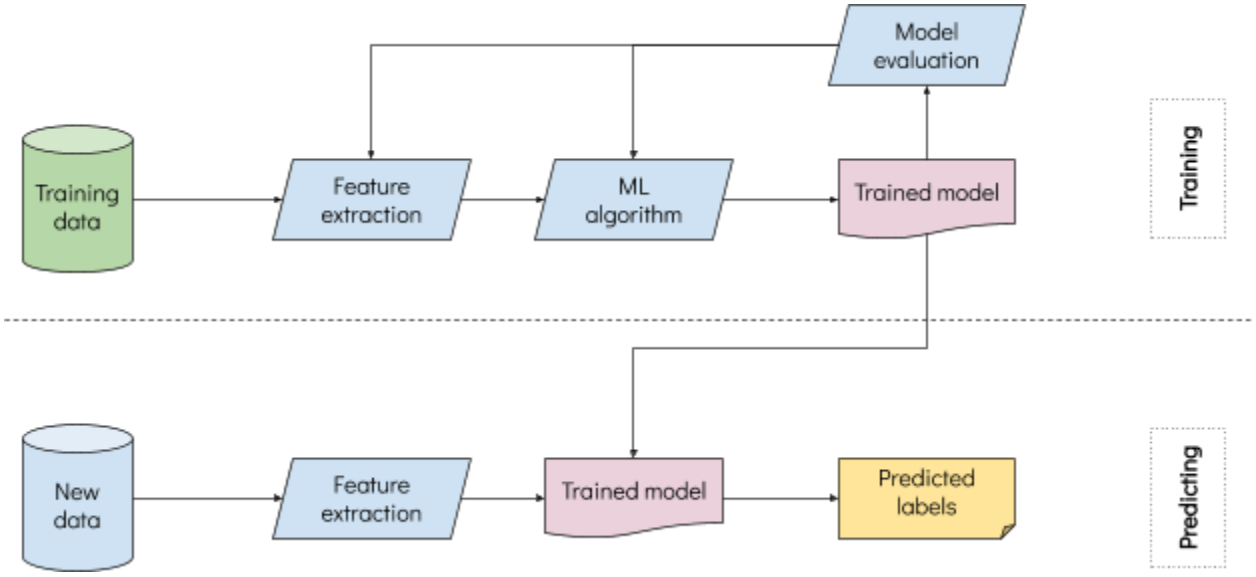


Figure S1: Flowchart of the Data Mining (DM) process.

The learning procedure of ML algorithms is divided into two broad categories, i.e. supervised and unsupervised learning, based on whether the output values (class) of the input samples are fed to the algorithm as prior knowledge or not (**Figure S2**). In the latter case the algorithm is expected to identify the underlying classes in the provided data.

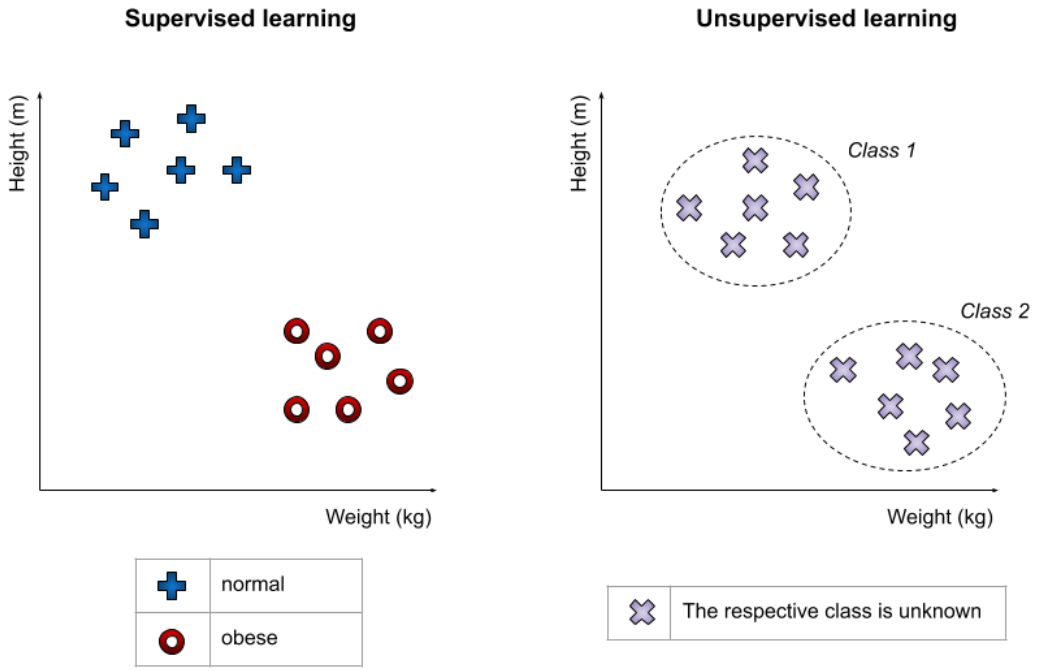


Figure S2: In supervised learning, the classes are already known and the algorithms aims to formulate a boundary that separates the given classes; in unsupervised learning the classes are unknown and the algorithm aims to “understand” the data and find inherent patterns or groupings.

Besides supervised and unsupervised learning there is another hybrid technique called semi-supervised learning which is often used when the unlabeled input data in a dataset are far more than the labeled ones. In semi-supervised learning the small amount of labeled input data is used as a starting point for training the algorithm, which is further trained with large amounts of unlabeled data. Supervised learning has two main branches, classification and regression; within a classification task the output values are a finite number of classes, whereas in the case of a regression problem the output variable is continuous. Unsupervised learning is largely represented by clustering where the algorithm aims to identify a set of clusters that are inherent to the input data (**Figure S3**).

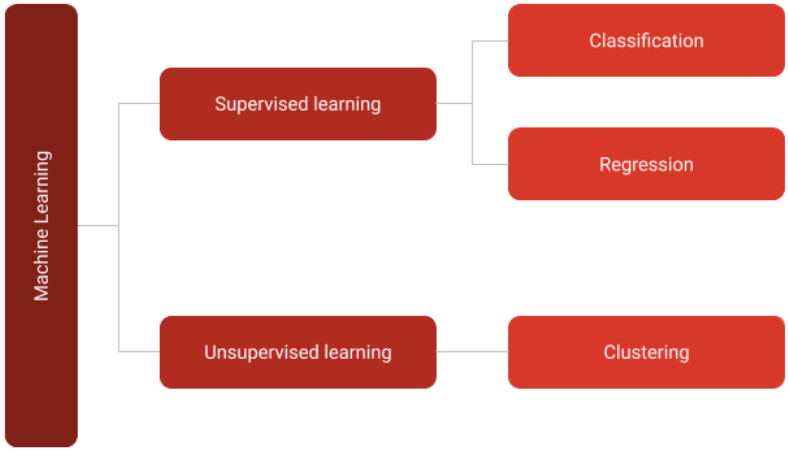


Figure S3: Supervised and unsupervised learning.

Overview of ML techniques

Over the past decades several machine learning algorithms have been presented in the literature, which differ in their approach, the type of data they input and output, and the type of task or problem that they are intended to solve. Below, we will describe briefly the most popular machine learning algorithms: Bayesian networks, Naive Bayes, Artificial Neural Networks, Decision Trees, Random forests and Support Vector Machines.

Bayesian Networks

A Bayesian network (belief network, directed acyclic graph model) is a model that is built based on the observed probabilistic relationship among a set of variables (e.g. symptoms and diseases); therefore its output is rather a probability than a prediction. Bayesian networks have been widely used in series of ML problems, including medical applications since they are able to provide reasoning for the reported outcomes as well as assign a probability representing confidence for each decision. As shown in **Figure S4** below, each node of the network is accompanied by a table of probabilities defined by the values of the variables it is connected to, i.e. the ones that affect its outcome. In the case that all employed variables are "naively" considered independent, the resulting algorithm is called Naive Bayes.

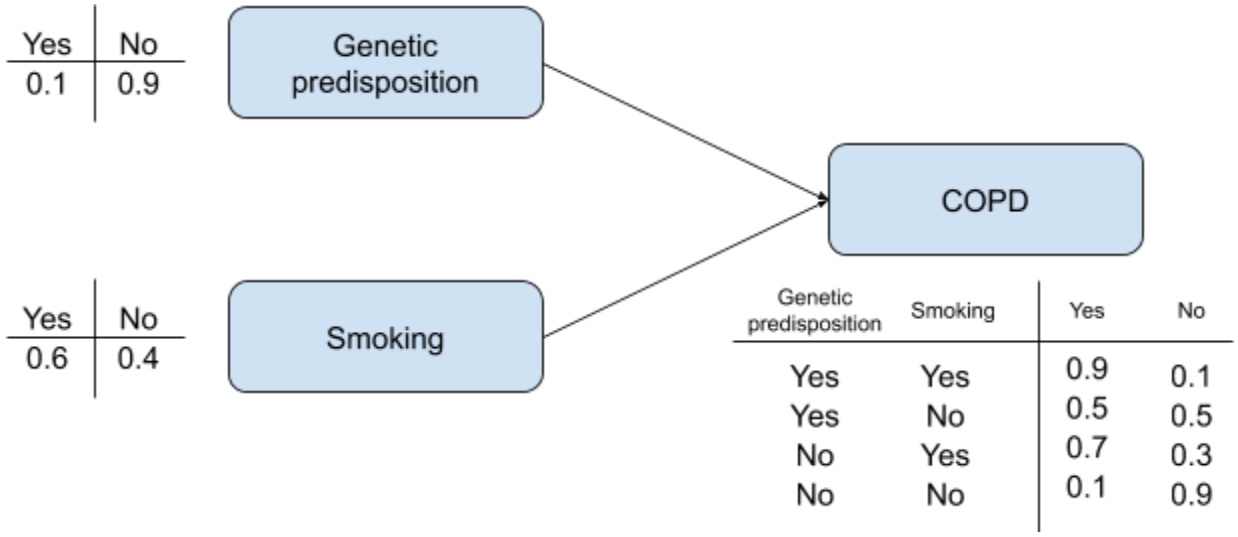


Figure S4: A provisional Bayesian network for COPD.

Artificial Neural Networks

Artificial Neural Networks are vaguely inspired by the notion and function of biological neural networks where neurons are interconnected by synapses and are trained to perform a specific task when activated. Artificial Neural Networks have proven quite useful in a series of tasks from various fields since they often perform very well. Due to their layered and often largely interconnected structure (**Figure S5**) the training process is quite time consuming and more importantly reasoning is almost impossible, therefore, they are often regarded as "black-boxes". Especially in medically oriented tasks this lack of explanation for the reported decision has attracted much criticism. Another concept that should be mentioned here is deep learning, that constitutes a subset of machine learning whereby the model resembles the layered approach of problem solving carried out by the human brain. Deep learning employs ANNs and a typical model often has at least three layers, where information is passed onto the next layer.

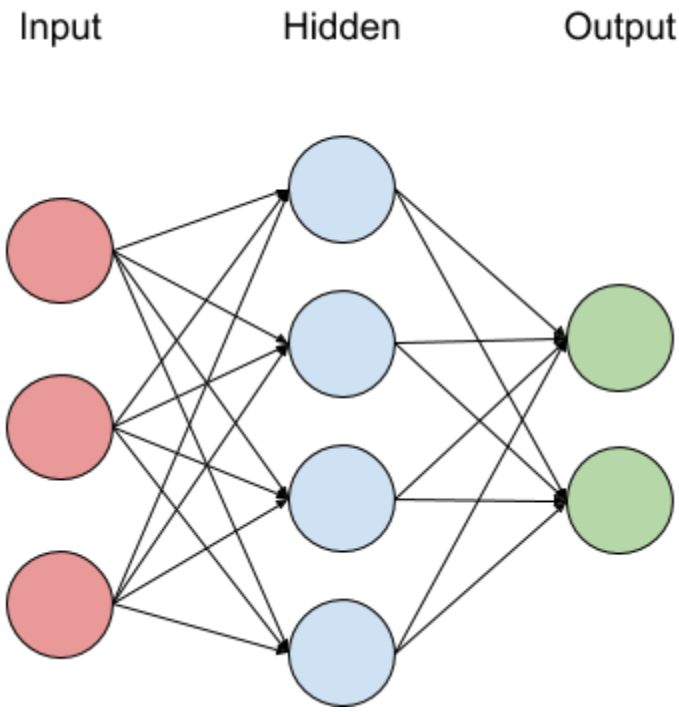


Figure S5: Architecture of an Artificial Neural Network with one hidden layer.

Decision Trees

Decision Trees constitute tree-structured classifiers where each node represents a variable and the leaves correspond to decision outcomes. The branches represent conjunctions of features that lead to the outcomes; by traversing the tree given the features values of a new sample, we are able to conjecture about its outcome. During the training phase where the tree architecture is formulated, the C4.5 algorithm is employed which often performs quite fast. The resulting architecture besides its simplicity, is also quite intuitive and transparent allowing for justified decisions. Specifically, each decision is based on a human-readable rule which provides adequate reasoning and subsequently makes Decision Trees a quite appealing solution for medical problems where transparency and reasoning are often prerequisites. **Figure S6** depicts a provisional architecture of a Decision Tree.

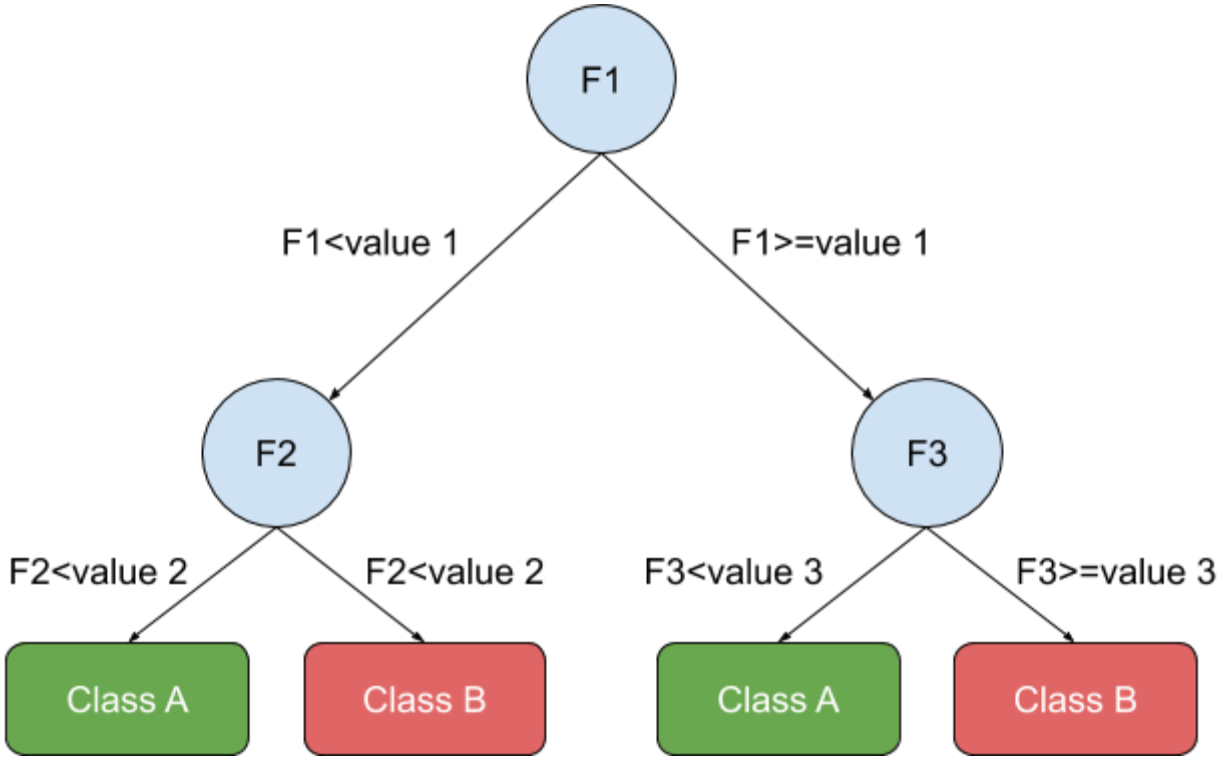


Figure S6: Provisional architecture of a Decision Tree classifier.

Random Forests

Random Forests or Random Decision Forests constitute an ensemble classifier that operates by constructing multiple Decision Trees in data subsets and assigning the output value by performing majority voting across the individual Decision Trees. **Figure S7** shows an exemplar Random Forest architecture.

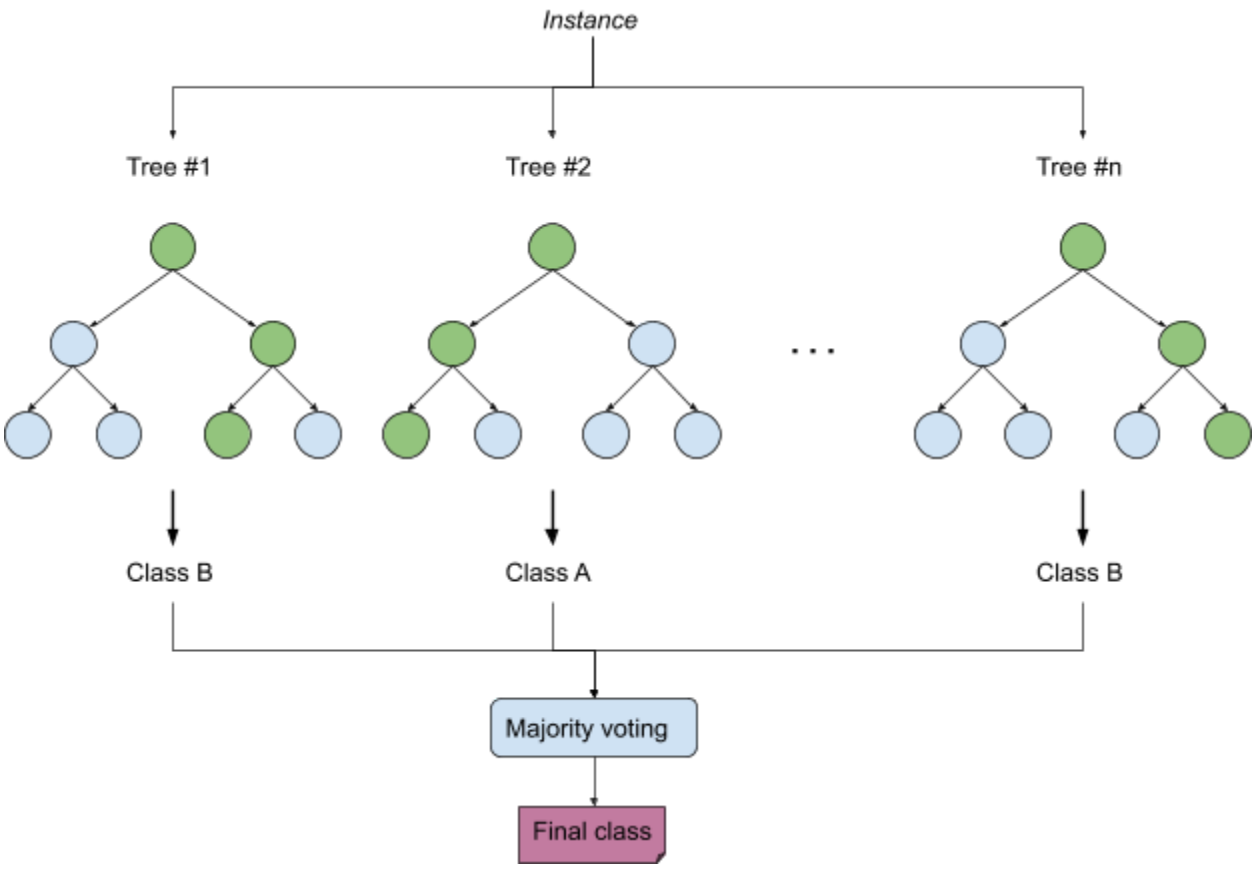


Figure S7: Architecture of a Random Forest algorithm.

Support Vector Machines

Support Vector Machines are one of the latest machine learning algorithms that has also been used extensively in medical and non-medical applications, due to the good performance and the generalization capability they often achieve. These two qualities are owed to the inherent process of training; specifically, Support Vector Machines map the initial input vector to a feature space of higher dimensionality where the samples can be separated with a linear hyperplane (kernel “trick”). Next, the algorithm searches across all possible hyperplanes that separate the samples in order to identify the one that maximizes the distance between the decision hyperplane and the most dubious instances (**Figure S8**).

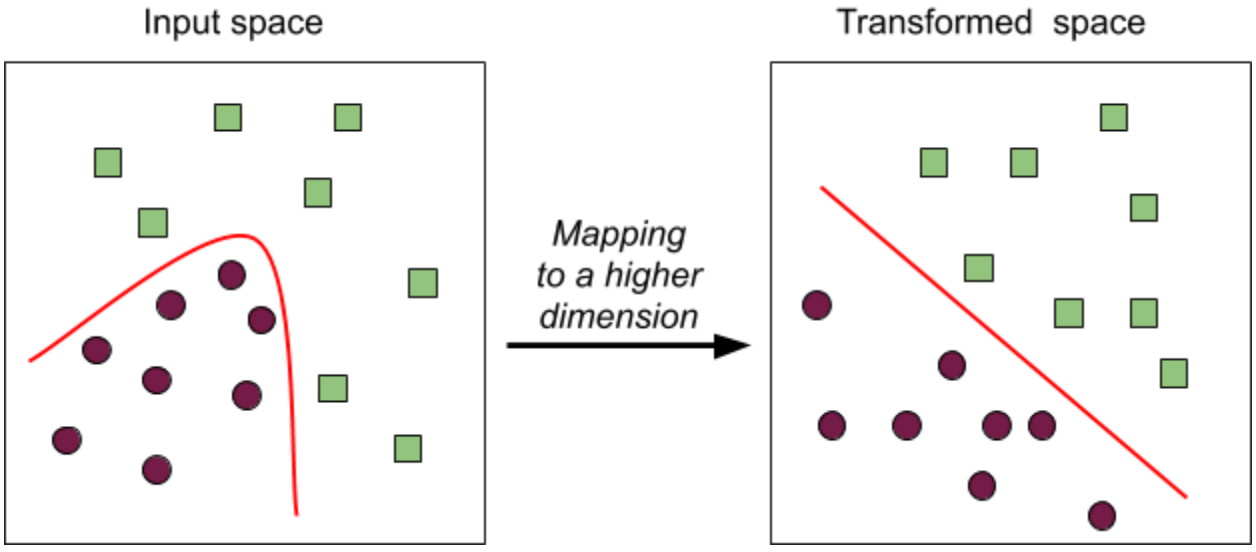


Figure S8: The kernel trick performed by the Support Vector Machines involves mapping the input vector to a higher dimensionality where the instances can be discriminated with a linear hyperplane.

Besides, the aforementioned algorithms, there are plenty of other machine learning algorithms, as well as variations of those algorithms with their respective strengths and limitations that heals towards deciding the most appropriate one for each task under consideration.

AI/ML validation

Within all “flavors” of AI or ML there are certain issues that need to be dealt with, that pertain to the fact that AI is essentially data-driven. When a model is trained with very limited data, these samples are memorized by the algorithm and the performance is nearly optimal for the specific dataset but very poor for other samples. This is much like a human that learns by heart a very specific task and is unable to perform well in other tasks. In a similar manner, an algorithm that is expected to discriminate between two classes and has been trained with an unbalanced dataset where one class is largely underrepresented, its performance towards disciminating that class will be relatively poor. This resembles a child that can recognize a basic set of common colors but when presented with one that has seen only a few times, it will most likely not recognize it.

All the aforementioned aspects regarding the performance of the algorithm are assessed quantitatively during the validation of the algorithm. For validation purposes the dataset is divided into two subsets, namely training and testing set where the latter is used in order to assess the performance of the trained model with new and previously unseen input data. Based on the size of the initial dataset, the testing set often contains 20%-40% of the input data. Another popular technique that is frequently used for validation purposes is *n*-fold cross

validation, whereby the initial dataset is partitioned in *n* equal subsets (or folds) from which *n-1* are used for training and the remaining one is used for testing; this process is repeated *n* times until all the folds have been used once for testing and the respective results are averaged in order to assess the overall performance of the model. A variation of *n*-fold cross validation is called Leave One Out Cross Validation (LOOCV) where *n* equals the total number of samples in the dataset. LOOCV is often indicated for limited datasets but is rather computationally intensive.

As for evaluation metrics, several ones have been described depending on the purpose of the machine learning algorithm, e.g. classification, regression, etc. The most widely used evaluation metrics are presented in **Table S2**.

Table S2: Most common metrics used for assessing the performance of ML algorithms.

Metric	**Formula**	**Description**
Sensitivity (Se) or True Positive Rate (TPR)	$TP/(TP+FN)$	Fraction of positive examples, predicted correctly by the model
Specificity (Sp) or True Negative Rate (TNR)	$TN/(TN+FP)$	Fraction of negative examples, predicted correctly by the model
Accuracy (Acc)	$TP+TN/(TP+FP+TN+FN)$	Overall correctness of the model, the ratio of correctly predicted outcomes and total number of examples
Receiver Operating Characteristic (ROC)	-	Graphical plot displaying the trade-off between the true positive rate and the false positive rate
Area Under ROC curve (AUC)	-	The two-dimensional area underneath the entire ROC curve
Positive predictive value (PPV)	$TP/(TP+FP)$	The proportion of positive results in the true positive results
Negative predictive value (NPV)	$TN/(TN+FN)$	The proportion of negative results in the true negative results
F1 score	$2*TP/(2*TP+FP+FN)$	The harmonic mean of PPV and Se

Kappa statistic	$[Pr(A) - Pr(E)]/[1 - Pr(E)]$ Pr(A): the percentage of observed agreement between the predictions and actual values Pr(E): the percentage of chance agreement between the predictions and actual values.	The agreement between the predicted results obtained by the model and the actual values

True Positive (TP): an outcome where the model correctly predicts the positive class.
True Negative (TN): an outcome where the model correctly predicts the negative class.
False Positive (FP): an outcome where the model incorrectly predicts the positive class.
False Negative (FN): an outcome where the model incorrectly predicts the negative class.

Artificial Intelligence and Asthma

As noted in the section 'Literature Review' of the main manuscript, the retrieved publications are divided into four categories, namely: (1) Asthma screening and diagnosis, (2) Patient classification, (3) Asthma management and monitoring, and (4) Asthma treatment. The articles from each category are summarized in a separate table where the respective studies can be compared by a set of qualitative and quantitative criteria or characteristics. In the first column ('*Ref*') we provide the reference for each study, the second column ('*ML algorithm*') shows the ML algorithm that was employed in the study. In cases where the study explored the performance of several ML algorithms, the best performing algorithm is reported. The third column ('*Sample size*') shows the total number of subjects or samples used in each study. The fourth column ('*Evaluation method*') shows the technique used for evaluating the performance of the proposed classification scheme; the fifth column ('*Performance*') contains a set of the most important reported metrics assessing the performance of the proposed work. In the last column ('*Important features*') we present the features reported in each study as being most important and informative.

Table S3, **Table S4** and **Table S5** contain studies related to 'Asthma screening and diagnosis', 'Patient classification' and 'Asthma management and monitoring', respectively.

Asthma screening and diagnosis

Table S3: Publications relevant to 'Asthma screening and diagnosis'.

Ref	ML algorithm	Sample size	Input features	Evaluation method	Performance	Important features
[14]	SVM	73	Capnography	LOOCV	Acc=94.52%, Se=97.67%, Sp=90%	Upward expiration (AR1), downward inspiration (AR2), sum of AR1 and AR2
[15]	SVM	60	Clinical (lung sound recordings)	LOOCV	Acc=93.3%	Exchange time of the instantaneous frequency
[16]	SVM	254	Clinical (medical record)	10-fold CV	Acc=98.59%,Se=98.59%,Sp=98.61%	
[17]	ANN & Fuzzy logic	780	Clinical (Portable spirometer)		Acc=97.32%	
[18]	HMM	16	Clinical (respiratory sounds)		Acc=94.91, Se=89.34%, Sp=96.28%	
[19]	k-NN	75	Forced oscillation technique parameters	N-fold CV, LOOCV	Se=82.9%, Sp=86.1%, AUC=0.91	Cross products of the FOT parameters: fr2, Xm.Cdyn [fr=resonance frequency, Xm=Mean respiratory reactance, Cdyn=Respiratory system dynamic compliance]
[20]	SVM	16	Clinical (phonopneumogr	LOOCV	Reliability (TPR*TNR)=97.36%	

			ams-respiratory sounds)			
[21]	ANN	112	Clinical (questionnaire, history)	10-fold CV	Acc=96.77%, Se=96.15%, Sp=100%	Wheezing episodes until 5th year, wheezing episodes between 3rd and 5th year, wheezing episodes until 3rd year, weight, waist's perimeter, seasonal symptoms, FEF25/75, number of family members, ICS
[22]	Fuzzy rules	278	Clinical		Se=88%, Sp=100%	
[23]	ANN		Clinical, epidemiological		AUC=0.903	
[24]	SVM	150 discharge summaries	Clinical (EMR)	10-fold CV	Acc=82%	
[25]	ANN	350	Clinical	CV		
[26]	LR	514	Clinical (EHR)	Training-Testing	Se=86%, Sp=98%	History of allergic rhinitis, eczema, family history of asthma, maternal history of smoking during pregnancy
[27]	ANN	254	Clinical	Training-Testing (70-30)	Acc=100%	Cough, symptoms of exercise induced asthma, humidity levels at home, emotional reactions, air pollution, wheeze, respiratory distress, hospitalization before 3 years of age, response to irritants, response to allergens, phlegm, allergies (both parents), pursiness
[28]	Fusion algorithm	170	Clinical (questionnaires)	10-fold CV	Se=98%, AUC=1	
[29]	SVM	30	Clinical (respiratory sounds)	Training-Testing	Acc=94.6%	
[30]	RF	132	Clinical	Training-Testing (80-20)	Precision=83%	Inhaler, MEF2575, Age, Smoker, Wheeze and Breath Shortness
[31]	Fuzzy rules & ANN	455	Clinical (spirometry, impulse oscillometry)	Independent test set	Acc=99%, Se=99%	
[32]	ANN	58	Clinical (breath sounds)	Independent test set	Se=94.6, Sp=100%	
[33]	ANN	48	Clinical (breath sounds)	Training-Testing (80-20)	Acc=92.8%	
[34]	ANN	827	Genomic (IgE reactivity)	Training-Testing (60-40)	Acc=78%	Allergens: Penicillin, Derm. Farinae, Kiwi, Timothy grass, Alpha amylase, Ph1 p1, Derp 1
[35]	ANN	51	Electronic nose, FeNO, and lung function testing	Training-Testing	Acc=95.8%	Electronic nose and FeNO
[36]	ANN	82	Genomic (SNPs)	5-fold CV	Acc=78%	
[37]	ANN	2832	Clinical (questionnaire)	Independent test set	PPV=100%	

[38]	ANN	10	Clinical (respiration sounds)	4-fold CV	Acc=80%	
[39]	ANN	180	Clinical (questionnaire)	Independent test set	Spearman rank order correlation coefficient=0.66	
[40]	SOM	32	Clinical (lung sounds)		Acc=78%, Se=52%	
[41]	DT	968	Clinical (lung function testing)	10-fold CV	PPV=66%, TPR=82%	
[42]	DT	12512	Clinical (spirometry, history, questionnaire, medication)	10-fold CV, Independent test set	Se=79%	
[43]	DT	26 signals	Clinical (lung sounds)	LOOCV	Acc=92%	
[44]	RF	554	Genetic (SNPs) and clinical	Bootstrapping	Acc=87%, AUC=0.84	Allergen sensitization, lung function markers
[45]	RF	461	Genetic and clinical	Training-Testing (80-20)	Se=97%, Sp=34%, AUC=0.82	Dust mite, pollens, pet allergens
[46]	GMM	24	Clinical (lung sounds)	LOOCV	Se=97.2%, Sp=94.2%, AUC=0.974	
[47]	LR	190	Genetic (nasal RNA)	Independent test set	AUC=0.994	
[48]	SVM	95 recordings	Clinical (respiratory sounds)		Acc=84%, Se=71.4%, Sp=88.9%	
[49]	DT	5032	Clinical (patient record)	5-fold CV	Definite asthma cases: PPV=66%,Se=98%,Sp =95%; Definite and probable asthma cases: PPV=82%, Se=96%, Sp=90%; Definite-probable and doubtful asthma cases: PPV=57%, Se=95%, Sp=67%	
[50]	SVM	283	Genetic (gene expression)	10-fold CV	Acc=95%	
[51]	RF	109	Exhaled breath condensate	Independent test set	Se=80%, Sp=75%	
[52]	RF	79	Genetic (micro RNA)	LOOCV	AUC=0.974	miR-125b, miR-16, miR-299-5p, miR-126, miR-206, miR-133b
[53]	ANN		Clinical	Independent test set	Acc=93%, Se=81%, Sp=100%	
[54]	k-NN	10	Clinical (lung sounds)	1-fold CV	Acc=77%	
[55]	ANN	60	Clinical	Training-Testing	Acc=43%	
[56]	JDINAC	461	Clinical	10-fold CV	Acc=86%, Se=84%, Sp=87%, AUC=0.94	Component-specific IgEs
[57]	LR & RF	177	Genomic (serum miRNA)	10-fold CV	Se=89%, Sp=77%, AUC=0.86	

[58]	NB	322	Clinical, patients history	10-fold CV	Acc=70.7%	
[59]	ANN		Capnogram		Acc=95.65%	
[60]	DT	1104	Clinical	10-fold CV	Se=93%, Sp=85%	Ever had asthma, current asthma, shortness of breath, atopy and wheezing, breathless but no family history
[61]	ANN & Fuzzy expert system	908	Genomic (SNPs)	Independent test set	Acc=94%	MS4A2 Glu237Gly, IL4Ra Glu375Ala

SVM: Support Vector Machine; ANN: Artificial Neural Networks; HMM: Hidden Markov Models; k-NN: k Nearest Neighbors; LR: Logistic Regression; RF: Random Forests; SOM: Self-organizing Maps; DT: Decision Trees; GMM: Gaussian Mixture Models; JDINAC: Joint density-based non-parametric differential interaction network analysis and classification; NB: Naive Bayes

Patient classification

Table S4: Publications relevant to 'Patient classification'.

Ref	ML algorithm	Sample size	Input features	Evaluation method	Performance	Important features
[67]	ANN	344	Genomic	5-fold CV	Acc=74.4%	
[68]	RF	96	Clinical	Training-Testing	Acc=70%, Se=81%, Sp=67%, AUC=0.86	15 VOCs
[54]	k-NN	10	Clinical (lung sounds)	1-fold CV	Acc=77%	
[55]	ANN	60	Clinical	Training-Testing	Acc=43%	
[56]	JDINAC	461	Clinical	10-fold CV	Acc=86%, Se=84%, Sp=87%, AUC=0.94	Component-specific IgEs
[57]	LR & RF	177	Genomic (serum miRNA)	10-fold CV	Se=89%, Sp=77%, AUC=0.86	
[58]	NB	322	Clinical, patients history	10-fold CV	Acc=70.7%	
[69]	Ensemble classifier	55	Clinical	LOOCV	PPV=95%	Tracheal wheeze sounds
[70]	Fuzzy Rules	28	Clinical (combination of 10 asthma severity scores)		Kappa coefficient=1	
[71]	DT	341	Clinical	10-fold CV	Se=84%, Sp=71%, AUC=0.83	
[72]	LASSO & stochastic gradient boosting	260	Clinical, Genomic	LOOCV	AUC=0.81	PKN2, PTK2, ALPP
[73]	SVM	346	Clinical	10-fold CV	Acc=81%, Se=62%, Sp=87%	-

[74]	DT	107	Clinical	10-fold CV	Acc=82.4%	Th2-mediated inflammation, corticosteroid insensitivity
[75]	GMM	1642	Clinical	CV	-	IL-13, IL-5
[76]	SVM	378	Clinical	LOOCV	Acc=93%	Age of asthma onset, quality of life, symptoms, medications, health care use
[77]	HMM	2255	Clinical	10-fold CV		Patterns of IgE responses over time
[78]	LR	1048	Clinical	10-fold CV	Acc=85%	
[79]	RF	348	Genomic	-	Misclassification rate=44%	ADAM33
[80]	DT	205	Genomic, Clinical	-	Acc=78%	Gene expression, clinical covariates, indicators of health outcomes
[81]	DT	3160	Clinical	Independent test set	AUC=0.72	Change in PEF, hospitalization for asthma, initial oxygen saturation on room air, initial PEF, risk stratification, emergency care of acute asthma
[59]	ANN		Capnogram		Acc=95.65%	
[60]	DT	1104	Clinical	10-fold CV	Se=93%, Sp=85%	Ever had asthma, current asthma, shortness of breath, atopy and wheezing, breathless but no family history
[82]	Fuzzy expert system	42	Clinical	-	Cohen kappa coefficient=1	
[83]	ANN	128	Clinical	10-fold CV	Acc=80%	
[84]	ANN	486	Clinical	Training-Testing	Acc=98.7%, Se=97.63%, Sp=97.83%	FEF25-75%
[85]	DT	872	Clinical	Independent test set	Cluster 1: Se=84.1%, Sp=96.3%; Cluster 2: Se=94.1%, Sp=99.5%, Cluster 3: Se=90.1%, Sp=99.3%; Cluster 4: Se=91.6%, Sp=91.9%	Comorbidities, adherence, cognitive dysfunction, depression
[86]	LR	12792	Patient records	Independent test set	AUC=0.67	Age, BMI, race, smoking history
[87]	BN	9801	Clinical	Independent test set	Average posterior probability=0.833	Eczema, wheeze, rhinitis
[88]	LR & SVM	1019	Clinical	5-fold CV	Short-term prediction=0.86; Long-term prediction=0.66	Obesity, allergy
[61]	ANN & Fuzzy expert system	908	Genomic (SNPs)	Independent test set	Acc=94%	MS4A2 Glu237Gly, IL4Ra Glu375Ala

SVM: Support Vector Machine; ANN: Artificial Neural Networks; JDINAC: Joint density-based non-parametric differential; interaction network analysis and classification; HMM: Hidden Markov Models; k-NN: k Nearest Neighbors; LR: Logistic Regression; RF: Random Forests; DT: Decision Trees; GMM: Gaussian Mixture Models; BN: Bayesian Networks

Asthma management and monitoring

Table S5: Publications relevant to 'Asthma management and monitoring'.

Ref	ML algorithm	Sample size	Input features	Evaluation method	Performance	Important features
[89]	Fuzzy expert system	25	Clinical (exacerbations)			
[90]	Ensemble classifier	96	Clinical, Patients record	5-fold CV	Acc=91.66%	Out of 140 initial variables,35 clinical variables were chosen
[91]	RF	42	Genomic	LOOCV	Acc=74%	20 features out of 30
[92]	RF	2	Clinical	10-fold CV	Acc=80.10%	FEV1, PEF,dust density, heart rate
[93]	Association rule mining	20959 ED visits	Environmental data, Patients records	Training-Testing	FDR=13%	SO2, NO, NO2, PM
[94]	Multiboost & Decision stumps	180	Clinical	10-fold CV	Acc=71.8%, Se=73.8%, Sp=71.4 %, AUC=0.757	
[95]	ANN & DT		Social media, Environmental data	10-fold CV	Precision=70%	asthma tweets, CO, NO2 and PM2.5
[96]	PCA & SVM	112	Clinical	10-fold CV	Se=95.54%	18 features
[97]	Pattern Based Decision Tree (PBDT) and Pattern Based Class-Association Rule (PBCAR)	33	Patient records, Clinical, Environmental data	Training-Testing (70-30)	PBCAR Acc=86.89%, Recall=84.12%; PBDT Acc=87.52%, Recall=85.59%	
[98]	ANN		Patients records, Clinical	CV	Acc=84%	
[99]	SVM	162	Clinical (cough signals)	-	Probability of correct classification=90%	-
[100]	RF	3206	Clinical, Patients records	Lasso penalization, out-of-bag estimation, CV, Ridge penalization	Critical care prediction: C-statistics=0.80, Se=79%; Hospitalization prediction: C-statistics=0.83, Se=75%	Advanced age, vital signs, arrival mode, comorbidities

[101]	RF	16	Clinical	LOOCV	Acc=87.4%, Se=47.2%, Sp=96.3%	Heart rate, respiratory parameters
[102]	ANN		Meteorological, Air pollution	CV	Acc=81%	
[103]	ANN	3602	Clinical, Meteorological, Air pollution	R2, Index of Agreement (IA), Root Mean Square Error (RMSE), Mean Bias Error (MBE)	0–4 years: R2=0.567; 5–14 years: R2= 0.207; 0–14 years: R2=0.528	
[104]	ANN	42	Clinical			
[105]	ANN		Clinical, Pollution data	Training-Testing	Acc=53%	Air pollution levels (NOx)
[106]	ANN	27	Clinical, Environmental data	CV		SO2, NO2, temperature, intake of medicine, relative humidity
[107]	Gradient boosting models	29354	Clinical, Patients records, Environmental, Air pollution, Neighborhood characteristics, Community viral load	3-fold CV	AUC=0.85	Oxygen saturation, pulse rate, respiratory rate, weight, age, triage acuity, weather variables
[108]	DT	200	Clinical	CV	Se=80%, Sp=89%	Dyspnea, accessory muscle use, wheezing
[109]	SVM	26	Clinical, Patient records (daily asthma diary)	Training-Testing	Acc=80%, Se=84%, Sp=80%	
[110]	RF	417	Clinical, Genomic	Independent test set	160-320 SNPs: AUC=0.66; 10 SNPs: AUC=0.57; Clinical traits: AUC=0.54	
[111]	Gradient boosting models	4548	Clinical, Environmental data	5-fold CV	AUC=0.78	Previous year bronchitic symptoms
[112]	XGBoost	7503	Air pollution, Meteorological data, Historical data	CV	AUC=0.832	Air pollution data, weather data, historical admissions data
[113]	LR	2691	Patients records	Training-Testing	Se=23%, PPV=56%, AUC=0.86	Number of ED visits in year 1, type of Insurance
[114]	BN	7001	Clinical	Training-Testing	Acc=100%, Se=100%, Sp=100%	63 variables out of 147 attributes
[115]	HMM		Clinical (respiration sounds)	CV	Se=85.7%	Cough
[116]	ANN & PCA	130	Clinical	3-fold CV	Se=100%, Sp=79.6%	FeNO, FEV1, FVC, FEV1/FVC, FEF25-75%

[67]	ANN	344	Genomic	5-fold CV	Acc=74.4%	
[68]	RF	96	Clinical	Training-Testing	Acc=70%, Se=81%, Sp=67%, AUC=0.86	15 VOCs
[58]	NB	322	Clinical, patients history	10-fold CV	Acc=70.7%	
[82]	Fuzzy expert system	42	Clinical	-	Cohen kappa coefficient=1	
[83]	ANN	128	Clinical	10-fold CV	Acc=80%	
[84]	ANN	486	Clinical	Training-Testing	Acc=98.7%, Se=97.63%, Sp=97.83%	FEF25-75%
[85]	DT	872	Clinical	Independent test set	Cluster 1: Se=84.1%, Sp=96.3%; Cluster 2: Se=94.1%, Sp=99.5%, Cluster 3: Se=90.1%, Sp=99.3%; Cluster 4: Se=91.6%, Sp=91.9%	Comorbidities, adherence, cognitive dysfunction, depression
[86]	LR	12792	Patient records	Independent test set	AUC=0.67	Age, BMI, race, smoking history
[87]	BN	9801	Clinical	Independent test set	Average posterior probability=0.833	Eczema, wheeze, rhinitis
[88]	LR & SVM	1019	Clinical	5-fold CV	Short-term prediction=0.86; Long-term prediction=0.66	Obesity, allergy

RF: Random Forests; ANN: Artificial Neural Networks; DT: Decision Tree; PCA: Principal Component Analysis; SVM: Support Vector Machine; LR: Logistic Regression; BN: Bayesian Network; HMM: Hidden Markov Model; NB: Naive Bayes