BIOCYBERNETICS AND BIOMEDICAL ENGINEERING 41 (2021) 1288–1302

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bbe

Original Research Article

A machine learning model for supporting symptom-based referral and diagnosis of bronchitis and pneumonia in limited resource settings

Katy Stokes [a], *Rossana Castaldo* [b,*], *Monica Franzese* [b], *Marco Salvatore* [b], *Giuseppe Fico* [c], *Lejla Gurbeta Pokvic* [d], *Almir Badnjevic* [e], *Leandro Pecchia* [a]

[a] *University of Warwick, Coventry, UK*
[b] *IRCCS SDN, Via E. Gianturco, Naples, Italy*
[c] *Life Supporting Technologies, Universidad Politécnica de Madrid, Madrid, Spain*
[d] *Medical Device Inspection Laboratory Verlab, Sarajevo, Bosnia and Herzegovina*
[e] *University of Sarajevo Sarajevo, Bosnia and Herzegovina*

ARTICLE INFO

Article history:
Received 8 January 2021
Received in revised form
9 August 2021
Accepted 6 September 2021
Available online 17 September 2021

Keywords:
Machine learning
Healthcare
Diagnostics
Screening tool
Pneumonia
LMICs

ABSTRACT

Pneumonia is a leading cause of mortality in limited resource settings (LRS), which are common in low- and middle-income countries (LMICs). Accurate referrals can reduce the devastating impact of pneumonia, especially in LRS. Discriminating pneumonia from other respiratory conditions based only on symptoms is a major challenge. Machine learning has shown promise in overcoming the diagnostic difficulties of pneumonia (i.e., low specificity of symptoms, lack of accessible diagnostic tests and varied clinical presentation). Many scientific papers are now focusing on deep-learning methods applied to clinical images, which is unaffordable for initial patient referral in LMICs. The current study used a dataset of 4500 patients (1500 patients affected by bronchitis, 3000 by pneumonia) from a middle-income country, containing information on subject population characteristics, symptoms and laboratory test results. Manual feature selection was performed, focusing on clinical symptoms that are easily measurable in LRS and in community settings. Three common machine learning methods were tested and compared: logistic regression; decision tree and support vector machine. Models were developed through a holdout process of training-validation and testing. We focused on six clinically relevant, easily interpreted patient symptoms as best indicators for pneumonia. Our final model was a decision tree, achieving an AUC of 93%, with the advantage of being fully intelligible and easily inter-

Abbreviations: GAPPD, Global Action Plan for the Prevention and Control of Pneumonia and Diarrhoea; CHWs, Community Health Workers; RDTs, Rapid Diagnostic Tests; LMICs, low- and middle-income countries'; LRS, limited resource settings

* Corresponding author at: IRCCS SDN, Via E. Gianturco, 113, 80143 Naples, Italy.
E-mail address: Rossana.castaldo@synlab.it (R. Castaldo).

https://doi.org/10.1016/j.bbe.2021.09.002

0168-8227/© 2021 The Authors. Published by Elsevier B.V. on behalf of Nalecz Institute of Biocybernetics and Biomedical Engineering of the Polish Academy of Sciences.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

preted. The performance achieved suggested that intelligible machine learning models can enhance symptom-based referral of pneumonia in LRS and in community settings.
© 2021 The Authors. Published by Elsevier B.V. on behalf of Nalecz Institute of Biocybernetics and Biomedical Engineering of the Polish Academy of Sciences. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Pneumonia has a devastating impact on global health and is the largest cause of death due to infection in children worldwide [1]. Although cases of pneumonia are found globally, the disease burden falls heavily in low- and middle income countries (LMICs), with around 90% of the global child mortality due to pneumonia and diarrhoea occurring in sub-Saharan Africa and South Asia [2]. Characteristics of individuals living in LMICs, such as malnutrition and exposure to air pollution, have long been understood to increase susceptibility to severe pneumonia [3]. Further to this, the access to appropriate treatment can be problematic in many LMICs, further increasing pneumonia fatalities. Fundamental to reducing deaths is rapid identification of the most appropriate treatment for pneumonia, as recognised in the 2015 Global Action Plan for the Prevention and Control of Pneumonia and Diarrhoea (GAPPD) [2], which calls for an end to preventable child deaths from pneumonia by 2025.

Once the patient is referred to specialised hospitals with suspected pneumonia, a confirmed diagnosis requires instrumental diagnostic tests (e.g., X-ray, pulse oximetry, blood tests and sputum tests), often integrated with more advanced investigations (e.g., CT scan, arterial blood gas tests, pleural fluid culture and bronchoscopy) [4,5]. Those tests and investigations are costly and not widely available in LMIC community settings, especially in rural areas. Unfortunately, pneumonia signs and symptoms are common to many other respiratory diseases (e.g., bronchitis), resulting in incorrect or delayed referrals. In fact, pneumonia signs and symptoms include cough (often with mucus and blood production), dyspnoea, fever, sweating, chest pain, loss of appetite, fatigue, nausea and vomiting (sometimes with mucus and blood) and confusion, especially in senior patients.

Accurate and timely pneumonia referral is crucial, especially in limited resource settings (LRS), which are abundant in LMICs [6]. The meaning of LRS may differ depending on the context [7], here LRS are taken to describe a healthcare setting experiencing a lack of either physical or organizational infrastructure, such as trained personal, facilities or equipment [6]. The COVID-19 pandemic demonstrated that appropriate referral of pneumonia is crucial also in high-income countries during a disaster, such as a pandemic [8]. For instance, the number of papers retrieved in PubMed combining the keywords “pneumonia” and “referral” moved from approximately 13 papers per month in 2019 to 90 papers per month in 2020 and 136 papers per month in 2021 until March (search query “*(referral) AND (pneumonia)*” in PubMed; search date 6th of April 2021).

The availability of large datasets and the need for highly accurate and timely referral and detection of diseases are motivating the use of data-driven machine learning (ML) methods in the field [9-11]. ML and deep learning methods have gained much attention in recent years for the automatic detection of pneumonia through imaging, in particular through analysis of chest X-ray or computed tomography (CT) [12-20]. The onset of COVID-19 and the subsequent global pressure on healthcare systems has further driven research in this area [21-24]. Such techniques are an attractive way to reduce the pressure on healthcare services with limited medical resources and staff, by providing fast and accurate diagnosis, reducing demand on equipment and expertise. Important considerations for use in LRS are speed of classification and minimal user input and energy requirements [12]. Although these methods are crucial for diagnosing pneumonia in a hospital setting, it is important to also consider that barriers exist in LMICs, which delay the diagnosis of pneumonia. In this regard, it is crucial to also support disease referrals in community settings, where symptoms alone can be assessed. In fact, it is widely accepted that identification and management of pneumonia in community settings significantly reduces deaths [25]. Healthcare services in LMICs strongly rely on community health workers (CHWs), especially in rural areas where there is inconsistent access to specialised doctors or hospitals [26]. Evidence suggested that Rapid Diagnostic Tests (RDTs) may support CHWs in detecting pneumonia in community settings [27,28]. Such an approach has proved promising in increasing quality of care and improving diagnosis and treatment availability especially in LRS [29], where patients may gain access to diagnostic systems using widely available technology such as smart phones [30]. Yet, there is a clear gap in interpreting pneumonia symptoms, which may be bridged through integration of RDTs and smart phones with ML, especially in community settings. This remains challenging due to the similarity of pneumonia symptoms with those of other respiratory diseases, such as bronchitis. In fact, different lower-respiratory-tract diseases tend to present with an overlapping set of symptoms, in this way it may be difficult to manually identify patterns and features in data, making it another appropriate challenge to be faced with ML. Unfortunately, this problem has not been well investigated. In fact, from the existing literature, it can be found that use of traditional ML for lower-respiratory-disease recognition based on symptoms is relatively scarce. This is limiting the potential for implementation of ML models in clinical analysis, which could improve the diagnostic capability of existing Computer-Aided Design (CAD) systems to automatically detect diseases such as pneumonia [31].

Overlapping symptoms represent a significant barrier for pneumonia referral, especially when resources and expertise are not widely available such as in LMICs and LRS. Correct referral holds the key to the most effective diagnosis, treatment and management [32]. In particular, pneumonia and bronchitis have many similar symptoms, affecting patients' referral. Indeed, where pneumonia refers to the presence of fluid in the alveoli, bronchitis is characterized by acute inflammation of the trachea and airways. Bronchitis most commonly results from viral infection, therefore, treatment with antibiotics is not generally effective, whereas for pneumonia the situation is reversed, providing strong motivation for correct referrals [17,33].

Several studies have achieved promising performance in classification of pneumonia [34-43] and other respiratory disease such as Chronic Obstructive Pulmonary Disease (COPD) and asthma [44-46] using ML with symptomatic predictors in combination with laboratory test results. However, there has been relatively little investigation of models using purely symptoms and signs as predictors [47,48], which, given the restraints upon healthcare at the community level, highlights a shortfall of research in this area. Furthermore, several issues in the field are apparent: reports do not provide a strong evidence base for their models, there is a lack of clarity in reporting of ML methods and the issue of distinguishing pneumonia from other similarly presenting respiratory diseases is not addressed in the existing literature.

Therefore, this study aimed to design an evidence based and interpretable ML model, using easily recognized symptoms and signs as predictors, which can distinguish between patients with bronchitis and pneumonia. Such a model is suitable for incorporation into a diagnostic tool, for the purpose of screening for pneumonia in the community, with the aim of improving access to referral and treatment.

2. Materials and methods

The steps completed during this work, beginning with the raw data and ending with an evaluation of predictive model performance, are outlined in Fig. 1. Symptomatic features were manually considered for machine learning in accordance with their clinical relevance and the goals of this study. A full description of the dataset and further details of each stage in the analysis are provided in the following sections.

2.1. *Dataset*

The data used in this work was collected as part of a prospective study, following internationally accepted medical practices for diagnosis of Chronic obstructive pulmonary disease (COPD) and asthma [49,50]. The dataset was generated to be suitable for design, validation, and real-time testing of a classifier to automatically identify bronchitis and pneumonia. Before starting the study, the ethics board approval for human subject testing from the Hospital Sarajevo was obtained (No. 01-11/EO-06/18), as well as the patients' informed consent.

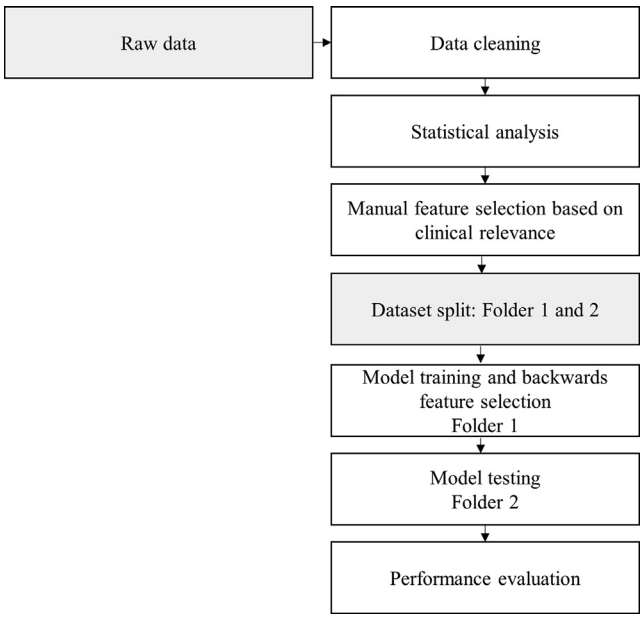

Fig. 1 – A workflow of all major steps taken in this work. Initially raw data was cleaned i.e., to deal with missing data and analyse outliers. Statistical analysis was performed to evaluate possible heterogeneity of population characteristics and variable distribution between patient groups. Initial manual feature choice on the clean dataset was completed before data splitting into separate folders. Feature selection was done by using backwards feature selection during the ML model training and validation. Model testing was performed on Folder 2. Finally, performance of the models on the test data folder was evaluated using several metrics recommended by the literature.

Healthcare institutions also approved all methods and procedures which were performed in accordance with the relevant guidelines and regulations. Samples originate from the period of October 2017 until December 2018.

Only patients with confirmed diagnosis were included as subjects in this study. Diagnoses were performed by medical professionals following clinical assessment according to international guidelines. Baseline assessments consisted of screening for patient symptoms using symptom-based questionnaires or interviews conducted by a medical professional. All spirometry lung function tests were obtained using the CareFusion "Master Screen" device (Hoechberg, Germany), which measured, derived and calculated all the required spirometry parameters.

The dataset comprised clinical information on 4500 individuals either diagnosed with bronchitis (1500) or pneumonia (3000). Information collected included a range of symptoms typical of respiratory illness, laboratory test results and various population descriptive characteristics such as exposure to air pollution or malnutrition. This information was established by medical professionals. Full description of the variables extracted are presented in Table 1.

Table 1 – Description of categorical and continuous variables present in dataset with different possible levels where appropriate for: multiclass and binary variables.

Sign and symptoms (Multiclass)	Description
CURB	Clinical prediction rule for mortality in CAP, risk of death after 30 days: 1: 2.7% 2: 6.8% 3: 14.0% 4: 27.8% 5: 27.8%
Auscultation	Examination: Listening to circulatory and respiratory systems: 1: Normal 2: Inspiration cracks 3: Pleural friction 4. Enhanced tune noise5: Unsuspecting noise 5: Unsuspecting noise
Sputum	On inspection: 1: Mucosal 2: Purulent 3: Haemoptysis
Results of X-ray (RTG)	1: Lobar 2: Segmental 3: Sub-segmental 4: Bronchopneumonia 5: In AIDS 6: Reverse 7: Non-segmental 8: No information
Sign and symptoms (binary)	
Associated diseases	Did the patient have any other associated diseases (yes or no)
Immunosuppression	Was immunosuppression present in patient (yes or no)
Allergy	Any allergies present in patient (yes or no)
Exposure to air pollution	Was the patient likely to be exposed to air pollution (probable or no)
Malnutrition	Did the patient suffer malnutrition
Cough	Did the patient have a cough (yes or no)
Expectoration	Did the patient have expectoration (yes or no)
Dyspnoea	Did the patient have dyspnoea (yes or no)
Pleura Pain	Did the patient have pleura pain (yes or no)
Fever	Did the patient have a fever (yes or no)
Sweating	Did the patient have sweating (yes or no)
Muscle pain	Did the patient have muscle pain (yes or no)
Headache	Did the patient have a headache (yes or no)
Loss of appetite	Did the patient have a loss of appetite (yes or no)
Laboratory test results	
Sedimentation	Measure of sedimentation of red blood cells.
Fibrinogen	Blood coagulation factor
CRP (mg/ml)	Indicative of inflammation or infection
Leukocytes	Total white blood cell count
Neutrophils	Subtype of white blood cell
Lymphocytes	Subtype of white blood cell
Monocytes	Subtype of white blood cell
Basophils	Subtype of white blood cell
Eosinophils	Subtype of white blood cell
Spirometry	Measure of amount and/or speed of air that can be inhaled and exhaled. Parameter measured: Volume expelled in one breath relative to expected reference, Forced Vital Capacity (FVC, %)

2.2. Statistical analysis

Statistical analysis was used to determine whether there was homogeneity between patients with pneumonia and bronchitis in terms of age, sex, and population characteristics. Further statistical tests were employed to understand the behavior of variables in terms of their distribution i.e., whether continuous variables were normally distributed. Finally, statistical differences of variables between the two patient groups were evaluated. Statistical test selection was informed by best practice guidelines described in the literature [51] and the methodology of the identified similar studies [34-42,52].

Features were tested for normality using the Chi-square goodness of fit test. Continuous variables were expressed as mean ± standard deviation or median and standard error. A non-parametric statistical test, Kruskal–Wallis test, was used for comparison of continuous variables between the two groups of patients (pneumonia and bronchitis), as it is appropriate for variables which are not normally distributed [51]. Categorical variables were expressed as a percentage and were compared using the Chi-Square, or Fisher Exact tests. A p-value of <0.05 was considered significant when assessing the variation of the features among the two patient groups. Bonferroni's correction was used for multiple hypothesis correction if necessary. Correlation analysis was carried out by Goodman and Kruskal's tau correlation.

Box and scatter plots were used to identify outliers in continuous variables. After ensuring that there were no changes to the data on importing or coding, outliers were quantified as any value which is more than three scaled absolute deviations from the median [53]. Subjects with outlying continuous variable values (156 individuals) were removed from the dataset, leaving 2844 pneumonia and 1500 bronchitis subjects. All the analyses were run in Matlab2019b.

2.3. Model training, validation, and testing procedure

As shown in Fig. 2, training and validation was performed on Folder 1 (60% of the total amount of data), and testing was performed on the remaining independent 40 % of data. The splitting was done in a stratified subject-wise fashion.

Feature Selection. Sufficiently large numbers of subjects allowed free selection amongst available attributes, complying with the '10 events per attribute' rule of thumb to avoid overfitting [9,54]. In the event that there were zero occurrences of a certain symptom in either class, these symptoms were discarded due to risk that information was not collected and to avoid a trivial separation of groups. Initial manual feature choice was performed on the clean dataset based on clinical relevance. In fact, for a diagnostic and/or screening application, the features should have some bearing on the disease [55]. Feature selection, based on the cluster of features manually selected, was then performed on Folder 1, after data splitting. Therefore, feature selection and model training were performed on the same folder [56]. Correlations between variables was evaluated using Kendall rank correlation test, with correlating variables not to be included as features. Correlated variables were considered as those which had a statistically significant Kendall rank coefficient greater than 0.5 ($|\tau| > 0.5,\ p < 0.05$). Backward feature selection was performed on the training dataset with only the best combination of features reported on.

Machine Learning Methods. Models automatically classifying patients as either having bronchitis or pneumonia were developed using three different machine learning methods: logistic regression (LR), decision tree and support vector machine (SVM). LR is an extension of linear regression, which predicts probability of a case belonging to a certain class [57]. A decision tree creates a set of 'if-else' conditions to predict the class of a given case [58]. SVM, which belongs to a general field of kernel-based machine learning methods, is used to efficiently classify both linearly and nonlinearly separable data [59]. Algorithm parameter tuning was performed during training and validation. Regarding the final model parameters, a fine tree with maximum splits of 100 was used for the decision tree, while a linear kernel with a scale of 0.8792 was used for SVM.

Training and validation. The training of the machine-learning models was performed on the folder 1 (1706 pneu-

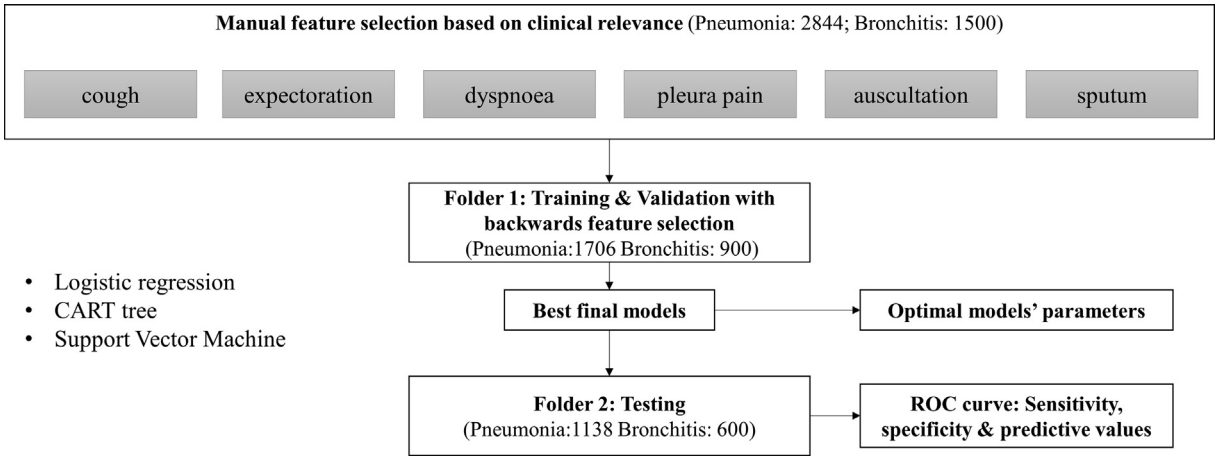

Fig. 2 – Feature selection, training and testing of machine learning methods. This diagram shows the workflow to develop the machine learning model.

monia patients, 900 bronchitis). Folder 1 was further divided into ten equal sized subsamples, according to the 10- fold person-independent cross-validation approach. Of these ten subsamples, nine subsamples were used as training data and the remaining one was retained for validating the model. The process was then repeated ten times, with each of the ten subsamples used exactly once as the validation data. Finally, the cross-validated estimations were computed by averaging the performances over the ten validation subsamples. Classification measures were adopted according to the standard formulae [60].

Testing. Testing a classifier involves analyzing its performance on a set of subjects that is independent from the training and validation set [61]. Accordingly, folder 2 (1138 pneumonia patients, 600 bronchitis) was used to test the trained models.

The model performance was obtained for the optimal operating point on the receiving operating characteristic (ROC) curve, as calculated by the MATLAB *perfcurve* function that relies on a previously described cost-function curve analysis [62].

Finally, the best performing model was selected as the one achieving the highest averaged area under curve (AUC), which is a reliable estimator of both sensitivity and specificity rates. In case of equal AUC, the model with the highest overall accuracy was selected.

3. Results

The clean dataset consisted of a total of 4344 samples, of which 2844 patients were diagnosed with pneumonia and the remaining with bronchitis. All continuous features were not normally distributed, with p-values <0.01. The mean, median, standard deviation and range of continuous variables is shown in Table 2. The final column reports the p-value of the Kruskal-Wallis Test for attribute variations between bronchitis and pneumonia subjects. All p-values fell <0.01, this indicates significant difference for all attributes between bronchitis and pneumonia.

The counts and proportions of the categorical variables between pneumonia and bronchitis groups is presented in Table 3. The final column reports the resulting p-value of the Chi-square (multi-class attributes) and Fisher Exact (binary attributes) Tests. Several symptoms were either not registered during data collection or not experienced by bronchitis sufferers: fever, sweating, muscle pain, headache and loss of appetite. Such attributes were discarded as the clear distinction does not provide an appropriate machine learning problem. Further, there may have been differences in data collection between groups for these symptoms. Age above 65 years old, auscultation, sputum and RTG showed to be statistically different with a p-value less than 0.01 between bronchitis and pneumonia cases.

Kendal rank correlation analysis between symptoms (cough, expectoration, dyspnoea, pleura pain, auscultation and sputum) and population descriptive variables (Above 65, associated chronic bronchopulmonary, immunosuppression, allergy, exposure to air pollution) found no correlations. Backwards feature selection found that including all the six above symptoms granted the best model performance. Population descriptive variables did not improve performance so were discarded in order to reduce the complexity of the model. Therefore, the final selected features were: cough, expectoration, dyspnoea, pleura pain, auscultation and sputum. Results of the three different ML methods are reported in Table 4. Although AUC are similar across the three methods, the model considered most successful and suitable was the decision tree. This is due to its superior overall accuracy over both LR and SVM. Furthermore, decision tree granted the fastest execution time to accurately predict pneumonia.

The ROC curves for the final models are shown in Fig. 3.

4. Discussion

This study proposes an easily interpreted, tree-based model for the automatic classification of pneumonia from bronchitis based entirely on easily measurable symptoms and signs. Features were selected based on their clinical relevance and availability on patient assessment. The manual feature selection method employed permitted a clear focus on the clinical utility and application of the model. Some key criteria used were: i) measurable in a point of care setting [63]; ii) parame-

Table 2 – Continuous variable statistics across bronchitis and pneumonia groups.

Variable	Bronchitis				Pneumonia				
	Mean	Median	SD	Range	Mean	Median	SD	Range	p-value
Age	53.04	52	27.08	93.00	38.29	36	23.28	100	<0.001
Sedimentation	15.39	15.62	8.70	29.92	57.78	58	25.11	85	<0.001
Fibrinogen	348.59	343.32	199.21	698.11	706.98	706	173.51	600	<0.001
CRP (mg/ml)	24.53	25.07	14.52	49.89	105.64	106	55.38	190	<0.001
Leukocytes	27.00	27.03	13.10	45.96	300.19	294.75	168.49	586.4	<0.001
Neutrophils	11.01	11.08	5.38	18.19	105.76	105.395	52.48	188.86	<0.001
Lymphocytes	5.42	5.34	2.67	9.20	53.15	53.355	27.46	95.45	<0.001
Monocytes	1.90	1.91	0.64	2.20	2.98	2.945	1.17	4	<0.001
Basophils	0.15	0.15	0.09	0.30	0.60	0.6	0.23	0.8	<0.001
Eosinophils	1.00	1	0.58	2.00	3.50	3.52	1.44	5	<0.001
Spirometry	65.34	65.485	8.92	30.96	57.62	57.87	10.09	34.98	<0.001

Mean, median, standard deviation (SD) and range are reported for pneumonia and bronchitis groups. The p-value reported corresponds to the outcome of the Kruskal–Wallis test for variation between pneumonia and bronchitis groups.

Table 3 – Categorical variable counts and percentages across pneumonia and bronchitis classes.

Variable	Bronchitis Count (%)	Pneumonia Count (%)	p-value
Sex (Female)	764 (50.93)	1398 (49.16)	0.28
Age Above 65	543 (36.20)	417 (14.66)	<0.001
Bronchopulmonary chronic Disease present	774 (51.60)	1425 (50.11)	0.35
Other Associated diseases	726 (48.40)	1476 (51.90)	0.03
Immunosuppression	741 (49.40)	1419 (49.89)	0.77
Allergy	769 (51.27)	1397 (49.12)	0.18
Probable exposure to air pollution	713 (47.53)	1440 (50.63)	0.06
Malnutrition	768 (51.20)	1416 (49.79)	0.39
Cough	775 (51.67)	1432 (50.35)	0.42
Expectoration	736 (49.07)	1405 (49.40)	0.85
Dyspnoea	765 (51.00)	1417 (49.82)	0.46
Pleura Pain	764 (50.93)	1428 (50.21)	0.66
Auscultation			<0.001
Auscultation 1	377 (25.13)	580 (580 20.39)	
Auscultation 2	370 (24.67)	527 (18.53)	
Auscultation 3	360 (24.00)	551 (19.37)	
Auscultation 4	393 (26.20)	592 (20.82)	
Auscultation 5	0 (0)	594 (20.89)	
Sputum			<0.001
Sputum 1	766 (51.07)	897 (31.54)	
Sputum 2	734 (48.93)	977 (34.35)	
Sputum 3	0 (0)	970 (34.11)	
RTG			<0.001
RTG 1	734 (48.93)	347 (12.20)	
RTG 2	766 (51.07)	359 (12.62)	
RTG 3	0 (0)	347 (12.20)	
RTG 4	0 (0)	385 (13.54)	
RTG 5	0 (0)	363 (12.76)	
RTG 6	0 (0)	357 (12.55)	
RTG 7	0 (0)	358 (12.59)	
RTG 8	0 (0)	328 (11.53)	

The p-value corresponds to the outcome of the chi-square or Fisher Exact tests for variation between pneumonia and bronchitis groups).

ters frequently investigated [64]; iii) ease of availability [65] and iv) reliability [34].

The results showed that by using a set of symptoms such as cough, expectoration, dyspnoea, pleura pain, auscultation and sputum, the model correctly identified more than 80% of patients with confirmed diagnosis of pneumonia. Although cough, expectoration, dyspnoea and pleura pain were not found to be statistically different among the two classes of respiratory diseases, the combination of those variables with auscultation and sputum signs achieved significant results to automatically distinguish patients affected by pneumonia from those with bronchitis. These results suggested that it was possible to correctly distinguish patients presenting with bronchitis or pneumonia, before performing clinical tests (e.g., X-ray), which required extensive expertise or advanced equipment. This can be crucial for patient referrals in community settings, especially in LRS. In fact, a set of predictors, which are easily recognised by CHWs or even self-reported and an automated ML model, which is suitable for incorporation into a tool such as an APP for use via mobile phones, would have great value in assisting referrals of pneumonia in LMICs. Moreover, in high income settings such a tool may complement traditional community healthcare services by providing widely available digital tests through apps for triage. The global need for such technology has been brought to the forefront of healthcare concerns in particular during outbreaks of COVID-19.

Although symptoms of bronchitis and pneumonia are similar, as discussed in the introduction, the required treatment is very different. First line treatment for pneumonia generally comprises antibiotic administration with close monitoring, while acute bronchitis is self-limiting and does not benefit from antibiotic treatment [35]. Therefore, there is a high cost to patient outcomes when these conditions are misdiagnosed or poorly referred, an easy-to-use tool capable of distinguishing the symptoms of these diseases is desirable.

The use of the symptoms identified in this study is largely supported by their identification as predictors of pneumonia by several authors in the literature, using different data and a variety feature selection methodology. The most commonly employed predictor from the literature is the outcome of auscultation, of which fast breathing specifically is strongly associated with pneumonia [34,36,40-42]. Cough [40,48] and productive cough/expectoration [36] were also found in several previous studies, as well as pleura pain [36,40] and dyspnoea [36,40,66]. Interestingly we were unable to identify any use of sputum evaluation in ML classifiers of pneumonia in the literature, which indeed in this study was found to be

TABLE 4 – Performance estimates on testing data.

	AUC	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Accuracy	F1 score	Execution time (s)
LR	93%	72%	88%	76%	86%	83%	87%	0.97
Decision Tree	93%	80%	84%	73%	89%	84%	87%	0.84
Linear SVM	93%	77%	85%	74%	88%	82%	85%	0.93

statistically different among bronchitis and pneumonia classes. This may reflect difference in clinical practices in different regions.

Three ML methods: SVM; decision tree and LR, were employed to facilitate comparison among commonly used models for clinical classification problems with varying interpretability. The selection was motivated by a desire to represent and compare methods which are frequently employed in the literature for similar problems, in particular between interpretable models. Comparison of performance between existing studies is limited for several reasons: variation in pneumonia reference standard; variation in subject population and lack of standardized reporting of ML methods and performance. The state-of-the-art studies to detect pneumonia and/or COVID-19 Pneumonia are reported in Table 5. Of the existing studies in the literature which utilized SVMs [35,39,67], only one specified distinguishing pneumonia from other diseases (as oppose to healthy patients). In this report from Rother et al. 2015 [67], the authors reported a program consisting of eight classifiers, with a sensitivity of 90% being the only performance parameter reported. Perhaps due to it is simplicity, LR has also proved a popular choice, Feng et al. 2020 [52] reported a high sensitivity of 100% with a specificity of 78% and AUC of 93% when identifying COVD-19 pneumonia based on symptoms and blood test results. Classification and Regression Trees (CARTs) were used by De Santis et al. 2017 [42] (sensitivity of 38%, specificity of 97%) and Steurer et al. 2011 [40] (AUC of 90%) to classify pneumonia, however the latter failed to use any internal or external validation techniques. Therefore, further analysis of the chosen methods in relation to detection of pneumonia was required. The results reported here support the use of simple, interpretable models such as LR and decision tree, which were shown to perform as well as or indeed better than linear SVM.

The tree-based model, which was considered most favourable for use in a clinical tool due to being easily understood, was found to give the best performance with an AUC of 93%. In comparison, in 2018 Pervaiz et al. found the WHO pneumonia symptomatic predictors for childhood pneumonia: cough; difficulty breathing, fever, tachypnoea and chest indrawing, to achieve an AUC of only 62% [36,47]. Reports of models built using symptoms/signs alone are relatively rare [42,67,68]. Nuzhat et al. reported high sensitivity and specificity (94%, 99%) in a logistic regression model using cough and lower chest wall indrawing as predictors, however the methodology lacked internal or external validation [68]. Rother et al. utilized an ensemble method, achieving a sensitivity of 90%, in this case with the disadvantage of an uninterpretable final model [67]. More common has been to include additional laboratory or clinical tests as predictors [34,37,40,41,69,70], however, such techniques are costly, time consuming and require specialist training, so would be unavailable in low resource settings. Other studies in the literature have focused on employing image-based classification to diagnosis pneumonia, in particular relating to detecting cases of COVID-19 (Table 5). When the patients are referred to hospital for suspected pneumonia, instrumental investigations are certainly needed, including X-ray imaging. In this case, image classification is most often approached through deep learning methods and achieved high

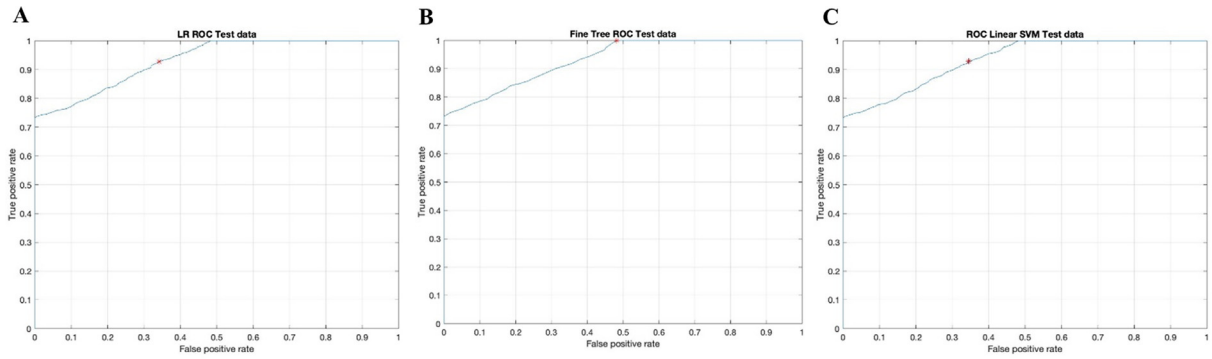

Fig. 3 – ROC curves for tested models. Panel (A) Logistic Regression; (B) Decision Tree; (C) Linear Support Vector Machine. All curves are reported based on test dataset which was held out from training.

performance; with the majority of methods reporting performance metrics above 90% (accuracy or AUC) [14,19-24]. For instance, Li et al. reported an AUC of 95% for detecting CAP from pulmonary CT scans [22], and Yue et al. achieved an accuracy of over 90% for pneumonia detection in five different convolutional neural networks (CNNs) using chest X-ray images [14]. Whereas, Wang et al. [16] and Stephen at al. [12] used a CNN-based model via chest X-ray images to detect pneumonia from other medical conditions and/or healthy patients by achieving higher performance than Sirazitdinov et al. [13], that used an ensemble of two convolutional neural networks for pneumonia localization from a large-scale chest X-ray database. Another study [19] by Nahid et al. employed a CNN model by also using chest X-ray images to detect patients affected by pneumonia achieving over 97% accuracy. A recent study by Musad et al. [18] employed radiomic features extracted from chest X-ray images via a CNN method which were then inputted to more traditional machine learning algorithms such as Random Forest Tree. However, they achieved a lower classification accuracy (86.3%) to discriminate among healthy, bacterial pneumonia and viral pneumonia categories. CNN was, in fact, often selected among the studies reported in Table 5 to automatically detect viral pneumonia via imaging techniques. Only one study by Srivastava et al. [20] applied CNN methodologies to assist medical experts by providing a detailed and rigorous analysis of the medical respiratory audio data for Chronic Obstructive Pulmonary detection.

In order to reduce the time and complexity of developing novel models from scratch, transfer learning has proved a promising fast route to building high performance deep learning models. In particular regarding rapid development of COVID-19 detection models, Hira et al. [21], Elgendi et al. [23] and Brunese et al. [24] achieved an accuracy of 97.5%, 94% and 97% respectively, in detecting pneumonia from chest X-rays using pre-trained models.

However, despite performing well on existing data, the deep learning methods used in these studies are less prone to be adopted in clinical settings due to low reliability and trustfulness. Moreover, the studies employing imaging techniques as predictors aimed to develop a diagnostic model, whereas in this study we aimed at developing a classifier that would have great value in assisting referrals of pneumonia, especially in LMICs. In fact, we presented a fully interpretable, tree-based model taking symptoms and signs as inputs, that can distinguish pneumonia patients with a similar performance (above 90%) to image-based deep learning approaches.

Nevertheless, the high performance seen from these very different ML approaches at both the initial patient referral (based on symptoms) and hospital confirmation (based on instrumental investigation through image analysis) highlighted the varied and exciting promise of AI for both referral and diagnosis of pneumonia, which may contribute to alleviating pressures on clinical staff and equipment especially in LRS of LMICs. Most importantly, early symptomatic discrimination of pneumonia from bronchitis may avoid unnecessary antibiotic treatments, helping to limit antibiotic resistance and avoiding the onset of pneumonia complications that may compromise patient treatment.

Among the studies aiming to develop tools for screening and/or diagnosis of respiratory disease such as pneumonia, bronchitis, asthma and COPD [44,45], the vast majority used additional laboratory test results such as white blood cell counts and sedimentation as well as symptoms. In practice, carrying out such tests requires a high level of expertise and costly facilities. Such requirements are not only challenging in LRS, but also take time, therefore are not an ideal basis for wide screening tools for rural areas in LMICs. Moreover, Pervaiz et al. found no benefit to adding oxyhaemoglobin levels to a ML model based on signs and symptoms alone to predict radiographically confirmed pneumonia in children. Furthermore, Naydenova et al. and Groeneveld et al. find that addition of patient C-Reactive Protein (CRP) levels to models based on symptoms, vital signs and age worsened performance in classification of pneumonia. This falls into a wider picture in which there is currently contention regarding the use of biomarkers such as CRP or Procalcitonin (PCT) as indicators for pneumonia [37,71], further research into their relevance, in particular in LMICs, is necessary to justify their use as predictors in diagnostic tools.

As well as demonstrating promising performance on existing data from a middle-income country, the model proposed has the advantage of being easily interpreted. Such explainable AI models have several benefits, which increases their clinical utility: trust in the system, guarding against bias, passing regulatory requirements, verifying outputs and

Table 5 – State-of-the-art studies to detect pneumonia.

Studies focused on signs and symptoms				
Author, Year	Final Predictors	Outcomes (Classes)	ML method and development (training, validation and/testing)	Performance estimates
Pervaiz et al., 2018 [47]	WHO criteria	Pneumonia; Acute respiratory illness	Logistic regression	AUC: 62%
Groeneveld et al., 2019 [75]	Runny nose absent; Feel ill; CRP concentration	Pneumonia; Healthy	Logistic regression	AUC: 75%
Nuzhat et al., 2017 [68]	Cough and lower chest wall indrawing combined	Pneumonia; Healthy	Logistic regression	SEN: 94% SPE: 99%
Naydenova et al., 2016 [76]	Respiratory rate; Heart rate; Temperature; Oxygen saturation; Age	Pneumonia; Healthy	Random Forest. Fivefold cross validation, data split for training and testing	ACC: 95.9% AUC: 99.7%
Grigull et al., 2012 [35]	14 clinical factors and vital signs including: age, temperature, blood pressure, etc.;12 laboratory parameters including: haemoglobin, leukocyte count, CRP level, etc.	Pneumonia; Other diseases	Voting algorithm using SVM, aNN and fuzzy logic	AUC: 99%
Steurer et al., 2011 [40]	Chronic cough; Daily fever; Dyspnoea; Respiratory rate; Pleural friction rub; CRP concentration	Pneumonia; No pneumonia	CART, leave-one-out cross validation	AUC: 90%
Rother et al., 2015 [77]	Included symptoms such as: Whistling/wheezing sounds and drowsy	Pneumonia; Other diseases	Program consisting of eight classifiers: SVM, ANN, fuzzy rule-based, random forest, LR, linear discriminant analysis, naïve Bayes, nearest neighbour, ensemble	SEN: 90%
van Vugt et al., 2013 [78]	Absence of runny nose; Breathlessness; Crackles; Diminishing breath sounds on auscultation; Tachycardia; Fever; CRP	Pneumonia; No pneumonia	Multilevel logistic regression, bootstrapping for internal validation	AUC: 77%
De Santis et al., 2017 [79]	Abnormal chest auscultation	Pneumonia; Other diseases	CART	SEN: 38% SPE: 97%
Feng et al., 2020 [69]	Combination of symptoms and signs (age, heart rate, fever, shiver, shortness of breath) and lab tests	COVID-19 pneumonia; Suspected COVID-19 pneumonia	Logistic regression (LASSO). Data split for training and testing	AUC: 93%

Table 5 – (continued)

Studies focused on signs and symptoms				
Author, Year	Final Predictors	Outcomes (Classes)	ML method and development (training, validation and/testing)	Performance estimates
DeLisle et al., 2013 [70]	Text from clinical notes and imaging report notes	Pneumonia; No pneumonia	Random fields probabilistic classifier	SEN: 58–75%
Studies focused on imaging				
Hira et al., 2020 [21]	Chest X-rays	COVID-19; Viral pneumonia; Bacterial pneumonia; Healthy	Deep transfer learning, pre-trained models. Nine convolutional neural network-based architectures: AlexNet, GoogleNet, ResNet-50, Se-ResNet-50, DenseNet121, Inception V4, Inception ResNet V2, ResNeXt-50, and Se-ResNeXt-50	ACC: 99.32% (binary) ACC: 97.55% (multi-class)
Li et al., 2020 [22]	Volumetric (3D) chest CT scans	COVID-19; Community-acquired pneumonia; Non-pneumonia abnormalities	COVNet deep learning framework: based on RestNet50. Training on 90%, testing on 10% of data	AUC: 95%
Brunese et al., 2020 [24]	Chest X-rays	COVID-19; Pneumonia; Pulmonary diseases; Healthy	Deep transfer learning, network based on VGG-16 (Visual Geometry Group model). Data split: training, testing and evaluation sets	ACC: 97%
Elgendi et al., 2020 [23]	Chest X-rays	COVID-19; Viral pneumonia; Bacterial pneumonia	Pre-existing neural network DarkNet-19. Data split: training 80%, validation 20%	ACC: 94.28%
Yue et al., 2020 [14]	Chest X-rays	Pneumonia; Healthy	Deep learning models, MobileNet (3×3 depthwise separable convolutions). Data split for training and testing	ACC: 92.9%
Sirazitdinov et al., 2019 [13]	Chest X-rays	Normal; No lung opacity/not Normal; Lung opacity	Ensemble of two convolutional neural networks, RetinaNet and Mask R-CNN. Data split for training and testing, training set further split (90% training, 10% validation)	SEN: 79.3% SPE: 75.8%
Wang et al., 2019 [16]	Chest X-rays	No pneumonia; Pneumonia; Other medical conditions	Deep learning, Cooperative CNN. Data split for training (95%) and testing (5%)	SEN: 89.04% SPE: 78.53%
Stephen et al., 2019 [12]	Chest X-rays	Pneumonia; Healthy	CNN. Data split for training, testing and validation	ACC: 93.73%
Walsh et al., 2018 [15]	CT scans	Fibrotic lung disease; Normal	Neural network Inception-ResNet-v2.16 algorithm. Data split for training, testing and validation	ACC: 76.4%

Table 5 – (continued)

Studies focused on signs and symptoms				
Author, Year	Final Predictors	Outcomes (Classes)	ML method and development (training, validation and/testing)	Performance estimates
Srivastava et al., 2021 [20]	Medical respiratory audio data	Normal; abnormal: COPD, asthma, lower and upper respiratory tract infection	CNN. Data split for training, testing and validation	ICBHI score: 93%
Masud et al., 2021 [18]	Chest X-rays	No disease; bacterial pneumonia; viral pneumonia	Deep learning to extract features, Random Forest classifier, training subsets via bootstrap sampling	ACC: 86.3%
Nahid et al., 2020 [19]	Chest X-rays	Pneumonia positive; pneumonia negative	CNN. Data split for training, testing and validation	ACC: 97.92%

CRP: C-Reactive Protein; SEN: Sensitivity; SPE: Specificity; ACC: Accuracy; AUC: Area under the ROC curve.

assessing risk [72]. Indeed, the importance of explainability in AI is not limited to symptom-based classifiers but extends also into the field of image analysis. This is challenging for deep learning, where it is not easy to follow the 'decision making' process leading to the final classification, unless specific tools for visualization of data significance are employed [73,74]. In fact, there is evidence in the literature of predictive deep learning models which are able to indicate the areas of a chest X-ray which contributed most to disease detection, which allows rapid identification of areas of interest to a radiologist in a hospital setting [24].

The work detailed here is a proof of concept that a simple, evidence-based ML model has the potential to perform well using symptoms and signs alone as predictors. However, this study comes with certain limitations. Firstly, the data driving this work was collected in a European middle-income country, therefore, may be of limited utility in populations of low-income countries (e.g., Sub-Saharan Africa), which indeed experience the greatest burden from pneumonia. In fact, pneumonia in Europe and pneumonia in Sub-Saharan Africa may have different causes (e.g., ageing and pollution vs Saharan desert dust), which may result in different symptoms. Furthermore, training of CHWs in low-income settings in auscultation, spirometry or sputum evaluation may not be easy, indeed equipment and expertise may not be available in emergency rooms even in high-income regions. Whilst the ML model reported in this study has the advantage of discriminating pneumonia from bronchitis, determining of the underlying pathogenesis and severity of pneumonia was beyond the scope of this study. This will be an interesting future avenue of research, as this information would be highly valuable in identifying the best treatment for pneumonia patients on an individual basis. Finally, in order to produce a clinical tool, it would be necessary to incorporate all commonly presenting respiratory diseases, not only bronchitis and pneumonia.

5. Conclusions

Correct referral and diagnosis of pneumonia is challenging due to low specificity of symptoms, lack of widely available diagnostic tests and varied clinical presentation amongst subpopulations. In this study, we applied machine learning algorithms to a dataset of 4344 patients (1500 bronchitis, 2844 pneumonia) containing information on subject population characteristics, symptoms and laboratory test results. Feature selection found 6 clinically relevant and easily interpreted patient symptoms to be the best predictors of pneumonia in this dataset. The best performing model that was able to distinguish pneumonia from bronchitis via sign and symptoms was a decision tree, which achieved an AUC of 93%. The robust, evidence-based design and ability to use symptoms to distinguish pneumonia from a similar respiratory disease (i.e., bronchitis) grants advantage for application in LMICs, compared to previously reported models relying mainly on instrumental tests and X-ray images. To be of most practical use in resource limited settings, machine learning models aiming at supporting disease screening, early diagnosis and appropriate referral, must provide thorough reporting of

methodology and performance and place emphasis on easily evaluated attributes such as presenting clinical signs and symptoms.

CRediT authorship contribution statement

Katy Stokes: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft, Writing - review & editing. **Rossana Castaldo:** Conceptualization, Methodology, Software, Formal analysis, Resources, Writing – original draft, Writing - review & editing, Supervision. **Monica Franzese:** Methodology, Writing - review & editing. **Marco Salvatore:** Methodology, Writing - review & editing. **Giuseppe Fico:** Writing - review & editing. **Lejla Gurbeta Pokvic:** Methodology, Investigation, Data curation, Writing - review & editing. **Almir Badnjevic:** Methodology, Investigation, Data curation, Writing - review & editing. **Leandro Pecchia:** Conceptualization, Methodology, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by EPSRC Impact Accelerator Award (EP/K503848/1 and EP/R511808/1). KS is funded by the MRC Doctoral Training Partnership [grant number MR/N014294/1]. RC, MS and MF are supported by "Progetti di Ricerca Corrente" funded by the Italian Ministry of Health.

REFERENCES

[1] Pneumonia. Fact Sheets: World Health Organization; 2019.
[2] Qazi S, Aboubaker S, MacLean R, Fontaine O, Mantel C, Goodman T, et al. Ending preventable child deaths from pneumonia and diarrhoea by 2025. Development of the integrated Global Action Plan for the Prevention and Control of Pneumonia and Diarrhoea. Arch Dis Childhood. 2015;100: S23.
[3] Berman S. Epidemiology of acute respiratory infections in children of developing countries. Rev Infect Dis 1991;13: S454–62.
[4] American Lung Association Scientific, Panel MER. Pneumonia symptoms and diagnosis. 2021.
[5] Tasaka S. Recent advances in the diagnosis and management of pneumocystis pneumonia. Tuberc Respir Dis (Seoul) 2020;83:132–40.
[6] Geiling J, Burkle Jr FM, Amundson D, Dominguez-Cherit G, Gomersall CD, Lim ML, et al. Resource-poor settings: infrastructure and capacity building. Chest 2014;146: e156S–e167.
[7] van Zyl C, Badenhorst M, Hanekom S, Heine M. Unravelling 'low-resource settings': a systematic scoping review with qualitative content analysis. BMJ Glob Health 2021;6 e005190.
[8] Pecchia L, Piaggio D, Maccaro A, Formisano C, Iadanza E. The inadequacy of regulatory frameworks in time of crisis and in low-resource settings: personal protective equipment and COVID-19. Health Technol 2020;10:1375–83.
[9] Foster KR, Koprowski R, Skufca JD. Machine learning, medical diagnosis, and biomedical engineering research - commentary. Biomed Eng Online 2014;13:94.
[10] Palaniappan R, Sundaraj K, Ahamed NU. Machine learning in lung sound analysis: a systematic review. Biocybernet Biomed Eng 2013;33:129–35.
[11] Wardlaw AJ, Rick EM, Pur Ozyigit L, Scadding A, Gaillard EA, Pashley CH. New perspectives in the diagnosis and management of allergic fungal airway disease. J Asthma Allergy 2021;14:557–73.
[12] Stephen O, Sain M, Maduh UJ, Jeong D-U. An efficient deep learning approach to pneumonia classification in healthcare. J Healthc Eng 2019;2019:4180949.
[13] Sirazitdinov I, Kholiavchenko M, Mustafaev T, Yixuan Y, Kuleev R, Ibragimov B. Deep neural network ensemble for pneumonia localization from a large-scale chest X-ray database. Comput Electr Eng 2019;78:388–99.
[14] Yue Z, Ma L, Zhang R. Comparison and validation of deep learning models for the diagnosis of pneumonia. Comput Intell Neurosci 2020;2020:8876798.
[15] Walsh SLF, Calandriello L, Silva M, Sverzellati N. Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study. Lancet Respir Med 2018;6:837–45.
[16] Wang K, Zhang X, Huang S, Chen F. Automatic detection of pneumonia in chest X-ray images using cooperative convolutional neural networks. 2019. pp. 328-40.
[17] Wu X, Song Z, Zhai X, Zuo L, Mei X, Xiang R, et al. Simultaneous and visual detection of infectious bronchitis virus and Newcastle disease virus by multiple LAMP and lateral flow dipstick. Poult Sci 2019;98:5401–11.
[18] Masud M, Bairagi AK, Nahid A-A, Sikder N, Rubaiee S, Ahmed A, et al. A pneumonia diagnosis scheme based on hybrid features extracted from chest radiographs using an ensemble learning algorithm. J Healthc Eng 2021;2021:8862089.
[19] Nahid A-A, Sikder N, Bairagi AK, Razzaque MA, Masud M, Z. Kouzani A, et al. A novel method to identify pneumonia through analyzing chest radiographs employing a multichannel convolutional neural network. Sensors. 2020;20.
[20] Srivastava A, Jain S, Miranda R, Patil S, Pandya S, Kotecha K. Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease. PeerJ Comput Sci 2021;7 e369.
[21] Hira S, Bai A, Hira S. An automatic approach based on CNN architecture to detect Covid-19 disease from chest X-ray images. Appl Intell 2020.
[22] Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, et al. Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy. Radiology 2020;296:E65–71.
[23] Elgendi M, Nasir MU, Tang Q, Fletcher RR, Howard N, Menon C, et al. The performance of deep neural networks in differentiating chest X-rays of COVID-19 patients from other bacterial and viral pneumonias. Front Med 2020;7.
[24] Brunese L, Mercaldo F, Reginelli A, Santone A. Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. Comput Methods Programs Biomed 2020;196 105608.
[25] Theodoratou E, Al-Jilaihawi S, Woodward F, Ferguson J, Jhass A, Balliet M, et al. The effect of case management on childhood pneumonia mortality in developing countries. Int J Epidemiol 2010;39(Suppl 1):i155–71.
[26] Marsh DR, Gilroy KE, Van de Weerdt R, Wansi E, Qazi S. Community case management of pneumonia: at a tipping point? Bull World Health Organ 2008;86:381–9.

[27] Elmardi KA, Malik EM, Abdelgadir T, Ali SH, Elsyed AH, Mudather MA, et al. Feasibility and acceptability of home-based management of malaria strategy adapted to Sudan's conditions using artemisinin-based combination therapy and rapid diagnostic test. Malar J 2009;8:39.
[28] Harvey SA, Jennings L, Chinyama M, Masaninga F, Mulholland K, Bell DR. Improving community health worker use of malaria rapid diagnostic tests in Zambia: package instructions, job aid and job aid-plus-training. Malar J 2008;7:160.
[29] Hamer DH, Brooks ET, Semrau K, Pilingana P, MacLeod WB, Siazeele K, et al. Quality and safety of integrated community case management of malaria using rapid diagnostic tests and pneumonia by community health workers. Pathog Glob Health 2012;106:32–9.
[30] 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;25:1213–7.
[31] Silva V, Novo ADR, Souza D, Rêgo A. Machine learning to assist in pneumonia decision making: A systematic review of the literature. SBC: Anais do VIII Symposium on Knowledge Discovery, Mining and Learning; 2020. p. 201–8.
[32] Badnjevic A, Cifrek M, Koruga D, Osmankovic D. Neuro-fuzzy classification of asthma and chronic obstructive pulmonary disease. BMC Med Inf Decis Making 2015;15:S1.
[33] Kinkade S, Long NA. Acute bronchitis.
[34] Naydenova E, Tsanas A, Howie S, Casals-Pascual C, De Vos M. The power of data mining in diagnosis of childhood pneumonia. J R Soc Interface 2016;13 20160266.
[35] Grigull L, Lechner WM. Supporting diagnostic decisions using hybrid and complementary data mining applications: a pilot study in the pediatric emergency department. Pediatr Res 2012;71:725–31.
[36] Haug PJ, Ferraro JP, Holmen J, Wu X, Mynam K, Ebert M, et al. An ontology-driven, diagnostic modeling system. J Am Med Inform Assoc 2013;20:e102–10.
[37] Groeneveld GH, van't Wout JW, Aarts NJ, van Rooden CJ, Verheij TJM, Cobbaert CM, et al. Prediction model for pneumonia in primary care patients with an acute respiratory tract infection: role of symptoms, signs, and biomarkers. BMC Infect Dis. 2019;19:976.
[38] Ge Y, Wang Q, Wang L, Wu H, Peng C, Wang J, et al. Predicting post-stroke pneumonia using deep neural network approaches. Int J Med Inf 2019;132 103986.
[39] Bejan CA, Xia F, Vanderwende L, Wurfel MM, Yetisgen-Yildiz M. Pneumonia identification using statistical feature selection. J Am Med Inform Assoc 2012;19:817–23.
[40] Steurer J, Held U, Spaar A, Bausch B, Zoller M, Hunziker R, et al. A decision aid to rule out pneumonia and reduce unnecessary prescriptions of antibiotics in primary care patients with cough and fever. BMC Med 2011;9:56.
[41] van Vugt SF, Broekhuizen BDL, Lammens C, Zuithoff NPA, de Jong PA, Coenen S, et al. Use of serum C reactive protein and procalcitonin concentrations in addition to symptoms and signs to predict pneumonia in patients presenting to primary care with acute cough: diagnostic study. BMJ 2013;346:f2450.
[42] De Santis O, Kilowoko M, Kyungu E, Sangu W, Cherpillod P, Kaiser L, et al. Predictive value of clinical and laboratory features for the main febrile diseases in children living in Tanzania: A prospective observational study. PLoS One. 2017;12:e0173314-e.
[43] Feng C, Huang Z, Wang L, Chen X, Zhai Y, Zhu F, et al. A novel triage tool of artificial intelligence assisted diagnosis aid system for suspected COVID-19 pneumonia in fever clinics. medRxiv. 2020:2020.03.19.20039099.
[44] Badnjevic A, Gurbeta L, Custovic E. An expert diagnostic system to automatically identify asthma and chronic obstructive pulmonary disease in clinical settings. Sci Rep. 2018;8:11645.
[45] Badnjevic A, Cifrek M, Koruga D, Osmankovic D. Neuro-fuzzy classification of asthma and chronic obstructive pulmonary disease. BMC medical informatics and decision making. 2015;15 Suppl 3:S1-S.
[46] Fraiwan L, Hassanin O, Fraiwan M, Khassawneh B, Ibnian AM, Alkhodari M. Automatic identification of respiratory diseases from stethoscopic lung sound signals using ensemble classifiers. Biocybernet Biomed Eng 2020;41:1–14.
[47] Pervaiz F, Chavez MA, Ellington LE, Grigsby M, Gilman RH, Miele CH, et al. Building a prediction model for radiographically confirmed pneumonia in Peruvian children: from symptoms to imaging. Chest 2018;154:1385–94.
[48] Nuzhat S, Ahmed T, Kawser CA, Khan AI, Islam SMR, Shahrin L, et al. Age specific fast breathing in under-five diarrheal children in an urban hospital: Acidosis or pneumonia? PLoS One. 2017;12:e0185414-e.
[49] Woodhead M, Blasi F, Ewig S, Garau J, Huchon G, Ieven M, et al. Guidelines for the management of adult lower respiratory tract infections-Full version. Clin Microbiol Infect 2011;17:E1–E59.
[50] Mandell LA, Wunderink RG, Anzueto A, Bartlett JG, Campbell GD, Dean NC, et al. Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. Clin Infect Dis 2007;44:S27–72.
[51] Parab S, Bhalerao S. Choosing statistical test. Int J Ayurveda Res 2010;1:187–91.
[52] Feng C, Huang Z, Wang L, Chen X, Zhai Y, Zhu F, et al. A novel triage tool of artificial intelligence assisted diagnosis aid system for suspected COVID-19 pneumonia in fever clinics. Ann Transl Med 2021;9.
[53] Aguinis H, Gottfredson RK, Joo H. Best-practice recommendations for defining, identifying, and handling outliers. Organ Res Methods 2013;16:270–301.
[54] Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373–9.
[55] Foster KR, Koprowski R, Skufca JD. Machine learning, medical diagnosis, and biomedical engineering research-commentary. Biomed Eng Online 2014;13:94.
[56] Castaldo R, Montesinos L, Melillo P, James C, Pecchia L. Ultra-short term HRV features as surrogates of short term HRV: a case study on mental stress detection in real life. BMC Med Inf Decis Making 2019;19:12.
[57] Hilbe JM. Logistic regression models. CRC Press; 2009.
[58] Lawrence RL, Wright A. Rule-based classification systems using classification and regression tree (CART) analysis. Photogramm Eng Remote Sens 2001;67:1137–42.
[59] Noble WS. What is a support vector machine? Nat Biotechnol 2006;24:1565–7.
[60] Parker C. An analysis of performance measures for binary classifiers. In: IEEE 11th International Conference on Data Mining. IEEE; 2011. p. 517–26.
[61] Foster KR, Koprowski R, Skufca JD. Machine learning, medical diagnosis, and biomedical engineering research - commentary. Biomed Eng Online 2014;13:94.
[62] Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 1993;39:561–77.
[63] Bartlett JG, Mundy LM. Community-acquired pneumonia. N Engl J Med 1995;333:1618–24.
[64] Kim ES, Park KU, Lee SH, Lee YJ, Park JS, Cho Y-J, et al. Comparison of viral infection in healthcare-associated pneumonia (HCAP) and community-acquired pneumonia (CAP). PLoS One. 2018;13:e0192893-e.

[65] Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med 2019;17:195.
[66] Porter P, Abeyratne U, Swarnkar V, Tan J, Ng TW, Brisbane JM, et al. A prospective multicentre study testing the diagnostic accuracy of an automated cough sound centred analytic system for the identification of common respiratory disorders in children. Respir Res 2019;20:81.
[67] 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-e.
[68] Nuzhat S, Ahmed T, Kawser CA, Khan AI, Islam SMR, Shahrin L, et al. Age specific fast breathing in under-five diarrheal children in an urban hospital: Acidosis or pneumonia? PLoS ONE 2017;12 e0185414.
[69] Feng C, Wang L, Chen X, Zhai Y, Zhu F, Chen H, et al. A novel triage tool of artificial intelligence-assisted diagnosis aid system for suspected COVID-19 pneumonia in fever clinics. medRxiv. 2021:2020.03.19.20039099.
[70] DeLisle S, Kim B, Deepak J, Siddiqui T, Gundlapalli A, Samore M, et al. Using the electronic medical record to identify community-acquired pneumonia: toward a replicable automated strategy. PLoS ONE 2013;8 e70944.
[71] Principi N, Esposito S. Biomarkers in pediatric community-acquired pneumonia. Int J Mol Sci 2017;18:447.
[72] Markus AF, Kors JA, Rijnbeek PR. The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies. J Biomed Inform 2020:103655.
[73] Porumb M, Iadanza E, Massaro S, Pecchia L. A convolutional neural network approach to detect congestive heart failure. Biomed Signal Process Control 2020;55 101597.
[74] Porumb M, Stranges S, Pescapè A, Pecchia L. Precision medicine and artificial intelligence: A pilot study on deep learning for hypoglycemic events detection based on ECG. Sci Rep 2020;10:1–16.
[75] Groeneveld GH, van't Wout JW, Aarts NJ, van Rooden CJ, Verheij TJM, Cobbaert CM, et al. Prediction model for pneumonia in primary care patients with an acute respiratory tract infection: role of symptoms, signs, and biomarkers. BMC Infect Dis. 2019;19:976.
[76] Naydenova E, Tsanas A, Howie S, Casals-Pascual C, De Vos M. The power of data mining in diagnosis of childhood pneumonia. J R Soc Interface 2016;13.
[77] Rother AK, 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.
[78] van Vugt SF, Broekhuizen BD, Lammens C, Zuithoff NP, de Jong PA, Coenen S, et al. Use of serum C reactive protein and procalcitonin concentrations in addition to symptoms and signs to predict pneumonia in patients presenting to primary care with acute cough: diagnostic study. BMJ 2013;346 f2450.
[79] De Santis O, Kilowoko M, Kyungu E, Sangu W, Cherpillod P, Kaiser L, et al. Predictive value of clinical and laboratory features for the main febrile diseases in children living in Tanzania: A prospective observational study. PLoS ONE 2017;12 e0173314.