



# The role of artificial intelligence in identifying asthma in pediatric inpatient setting

Gang Yu<sup>1</sup>, Zheming Li<sup>1</sup>, Shuxian Li<sup>2</sup>, Jingling Liu<sup>2</sup>, Moyuan Sun<sup>3</sup>, Xiaoqing Liu<sup>3</sup>, Fenglei Sun<sup>3</sup>, Jie Zheng<sup>3</sup>, Yiming Li<sup>3</sup>, Yizhou Yu<sup>3</sup>, Qiang Shu<sup>4</sup>, Yingshuo Wang<sup>2,4</sup>

<sup>1</sup>Department of IT Center, The Children's Hospital, Zhejiang University School of Medicine, Hangzhou, China; <sup>2</sup>Department of Pulmonology, The Children's Hospital, Zhejiang University School of Medicine, Hangzhou, China; <sup>3</sup>DeepWise AI Lab, Beijing, China; <sup>4</sup>National Clinical Research Center for Child Health, Hangzhou, China

*Contributions:* (I) Conception and design: Y Wang, M Sun, X Liu, Y Yu; (II) Administrative support: Q Shu, G Yu, Z Li, S Li, J Liu, J Zheng, Y Li, Y Yu; (III) Provision of study materials or patients: Y Wang, G Yu, Q Shu; (IV) Collection and assembly of data: Y Wang, Q Shu, G Yu, S Li, J Liu, F Sun, J Zheng; (V) Data analysis and interpretation: Y Wang, M Sun, X Liu, Y Yu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

*Correspondence to:* Qiang Shu. National Clinical Research Center for Child Health, No.3333 Binsheng Road, Hangzhou, China.

Email: shuqiang@zju.edu.cn; Yingshuo Wang. Department of Pulmonology, The Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, No. 3333 Binsheng Road, Hangzhou, China. Email: wangyingshuo@zju.edu.cn.

**Background:** The incidence of asthma in Chinese children has rapidly increased as a result of inadequate management. This is mainly due to the failure of many primary-level pediatricians to distinguish asthma from common respiratory diseases, such as bronchitis and pneumonia. Such misdiagnoses often lead to the abuse of antibiotics and systemic glucocorticoids. Additionally, if asthma is not diagnosed early, chronic airway inflammation results in lesions that not only hamper children's athletic abilities, but serve as the primary cause for adult chronic airway diseases, such as chronic obstructive pulmonary disease (COPD).

**Methods:** A number of machine learning-based models including CatBoost, Logistic Regression, Naïve Bayes, and Support Vector Machines (SVM) have been developed to identify asthma via utilizing retrospective electronic medical records (EMRs) of patients. These models were evaluated independently using EMRs from both the Pulmonology Department and other departments of the Children's Hospital, Zhejiang University School of Medicine, China.

**Results:** Two independent test sets were applied for performance evaluation. TestSet-1 consisted of 325 positive asthma cases and 428 negative cases from the Pulmonology Department. TestSet-2 was composed of 2,123 cases from non-pulmonology departments, and included 337 positive and 1,786 negative cases. Experimental results showed that the CatBoost model outperformed other models on both test sets with an accuracy of 84.7% and an area under the curve (AUC) of 90.9% on TestSet-1, and an accuracy of 96.7% and an AUC of 98.1% on TestSet-2.

**Conclusions:** The artificial intelligence (AI) model could rapidly and accurately identify asthma in general medical wards of children, and may aid primary pediatricians in the correct diagnoses of asthma. It possesses great clinical value and practical significance in improving the control rate of asthma in children, optimizing medical resources, and limiting the abuse of antibiotics and systemic glucocorticoids.

**Keywords:** Pediatric; asthma; artificial intelligence (AI); machine learning; diagnostic assistant

Submitted Mar 14, 2020. Accepted for publication Sep 18, 2020.

doi: 10.21037/atm-20-2501a

View this article at: <http://dx.doi.org/10.21037/atm-20-2501a>

## Introduction

Asthma is pathologically characterized by chronic airway inflammation. It is estimated that over 10 million children under the age of 14 have asthma. Furthermore, its incidence among Chinese children is rising rapidly with a relatively poor control rate (1-3). This is in part attributable to the failure of primary pediatricians to distinguish asthma from common respiratory tract virus infections in children, with asthma often being misdiagnosed as bronchitis or pneumonia. The average accuracy of asthma diagnosis among pediatricians at all levels was reported to be only 59.29%, and the level of diagnosis in primary units was significantly lower than that of tertiary class hospitals (4). Poor asthma control results in excessive and inefficient use of medical resources and leads to the abuse of antibiotics and systemic glucocorticoids. Moreover, this not only develops the continuous progression of chronic inflammatory diseases of airways that hamper the athletic abilities of children, but is also a significant cause of chronic airway diseases, such as chronic obstructive pulmonary disease (COPD), in adults (5). Therefore, an automatic asthma detection mechanism for children is critical for the efficient and accurate diagnoses of asthma by primary pediatricians.

With recent advancements in machine learning technologies, researchers have conducted novel studies on asthma. Latent class analysis can be used to identify a set of heterogeneous diseases with the diagnostic label of asthma (6), but such approaches cannot be directly applied to diagnosing asthma from the electronic medical records (EMRs) of patients. Prosperi and Marinho (7) used logistic regression and random forests to identify asthma, wheeze, and eczema. However, their model used a large set of attributes including 223 non-genetic variables and 215 single-nucleotide polymorphisms. The need for such large variables will not only lower the efficiency of the model, but also reduce its applicability in real world settings since generating such a large set of variables is not common in practice. Additionally, the data set used was small with only 554 subjects, which might have been insufficient in obtaining a general model. Tomita *et al.* (8) explored the effectiveness of deep neural networks in modeling combinations of symptom-physical signs and objective tests, using a total of 22 variables to predict the initial diagnosis of adult asthma. They used fewer variables than the study reported by Prosperi. However, deep neural networks-based methods are more sensitive to missing values commonly

observed in real clinical practice, thereby making their model less practical. A dataset with only 566 EMRs was another limitation of their work.

The present study thus aimed to identify an effective and efficient artificial intelligence (AI) model which can be used to assist pediatricians in diagnosing asthma in real clinical settings. It is hoped that such a model can eventually be deployed as a personalized diagnostic tool to lower the misdiagnosis rate of asthma among children and to reduce the misuse of antibiotics and systemic glucocorticoids.

## Methods

Study approval was granted by The Institutional Review Board (IRB) of the Medical Ethics Committee of Children's Hospital, Zhejiang University School of Medicine, China (IRB approval ID: 2020-IRB-039). The procedures were performed in accordance with the Declaration of Helsinki (as revised in 2013) and relevant guidelines & regulations. Informed consent was obtained after the procedure was fully explained to all participants and their legal guardians.

## Datasets

In this study, two sets of retrospective EMRs of patients under the age of 14 were collected from the Children's Hospital, Zhejiang University School of Medicine, China. DataSet-1 consisted of 3,761 cases, with 1,624 positive asthma cases and 2,137 negative cases from the Pulmonology Department. DataSet-2 was composed of 2,123 cases with 337 positives and 1,786 negatives from non-pulmonology departments including the Cardiovascular Department, Endocrinology Department, Nephrology Department, Neurology Department, and Hematology Department of the hospital. All records were reviewed by at least two respiratory experts who performed independent asthma diagnoses for each record, based on the guide of Children's bronchial asthma diagnosis and prevention (2016 version) (9). Initially, two respiratory experts issued their independent assessment. Subsequently, the gold standard for diagnostic decision (i.e., ground truth label for model training) for each record was made by judgment consistency. Where a discrepancy between the two independent judgments was apparent, additional independent assessments were performed by another two respiratory experts. The second round of expert opinions prevailed if a diagnostic agreement was made. In cases where an agreement was not achieved by the second review,

**Table 1** Raw EMR data in English translation

Fields	Descriptions
Gender	Male
Birth date	2011-07-17 11:13:00
Chief complaint	Cough, wheezing for 6 days
Physical examinations	Body temperature was 36.5°C, heart rate was 130 beats/min, breathing was 36 beats/min, blood pressure was 112/60 mmHg, and percutaneous oxygen saturation was 96%. Clear, spirited, superficial lymph nodes are not enlarged, ecchymosis on the back (after scraping), no cyanosis around the mouth, negative trident sign, pharyngeal congestion, thick lung sounds, wheezing sound can be heard and phlegm sounds, heart rhythm, no obvious pathological murmur, soft abdomen, liver and spleen under the ribs and swelling, neurological examination was normal, no obvious rashes were seen throughout the body
Lab tests	Blood routine (2014.12.25 in our hospital) white blood cell count $8.02 \times 10^9/L$ , neutrophil 60.1%, hemoglobin 129 g/L, platelet count $436 \times 10^9/L$ , CRP 4mg/L; chest radiograph (2014.12.25 our hospital): pneumonia, with a small amount of mediastinal emphysema, interstitial emphysema of the lower left lung may be
History of current disease	The child had cough, wheezing, paroxysmal continuous cough at home 6 days ago without obvious incentives, each time coughing 5-6 sounds, sputum is not easy to cough out, redness, chest tightness, and occasional abdominal pain during cough drama, wheezing is obvious at night and when the activity is intense, there is shortness of breath, and can be relieved spontaneously, with fever at the beginning of the disease, body temperature of 37.5 °C, no fear of chills, chills, no cyanosis, no barking cough, no hemoptysis, no rash, no nausea, Vomiting, no bloating, diarrhea, no night sweats, the next day she was treated at the “Lin’an People’s Hospital”, given “cefuroxime” intravenous drip treatment for 3 days (12.22–12.24), “methadone” (12.23) intravenous drip treatment for 1 day After treatments such as asthma and scraping, the child’s cough and asthma did not improve significantly, so he came to the hospital for medical examination, blood test (12.25 in our hospital) white blood cell count $8.02 \times 10^9/L$ , neutrophil 60.1%, hemoglobin 129 g/L, platelet count $436 \times 10^9/L$ , CRP 4 mg/L, chest radiograph (12.25 in our hospital): pneumonia, with a small amount of mediastinal emphysema, the left lower lung interstitial emphysema may be given, “Xi Shumei, methylprednisolone 30 mg QD” static After 2 days of instillation and nebulization treatment (12.25–12.26), the child’s condition improved slightly. In order to seek further treatment, “pneumonia” is planned to be admitted to the hospital. The sick child had a good spirit, poor appetite, good sleep, no obvious abnormalities in urine and urine, no significant increase or decrease in body weight, and denied the history of foreign body inhalation and choking.
History of diseases	The child had previous “cough and wheezing 5 times” and was about 3 years old at the first onset. Denied history of convulsions. Denied a history of allergies such as food. Denied a history of infectious diseases. In May 2014, he was diagnosed as “bronchial asthma” in “Lin’an People’s Hospital” without formal treatment
Family history of diseases	Denied the history of genetic diseases and infectious diseases in the family

the process was repeated until a diagnostic agreement was reached.

### Feature extraction

As shown in *Table 1*, the collected raw dataset consists of free texts arranged in a set of predefined fields in EMRs such as gender, birth date, chief complaint, physical examinations, lab tests, history of current disease, family history of diseases, history of diseases, and so on.

In order to train machine learning-based models, features were extracted from the relevant raw EMR fields.

Due to the semi-structured property of the raw EMR data, traditional rule-based nature language processing (NLP) methods were utilized for feature extraction. In this study, a set of NLP methods including regular expression, word distance, synonym analysis, etc. were applied to convert the raw free texts into features with either numerical or yes/no (1/0 binary) values. For example, the regular expression for total serum immunoglobulin E (IgE) values were written as “IgE [[:><]]\*?([+]| [0-9]+\+ )”. For all binary features (i.e., yes/no), zero was the default value, unless the corresponding regular expression identified given key words with no surrounding negatives in a thesaurus; in this case,

**Table 2** Extracted features and their properties

EMR fields	Extracted features	Property/values
Chief complaints	Cough	Yes/No (1/0)
	Wheezing	Yes/No (1/0)
	Tachypnea	Yes/No (1/0)
	Chest tightness	Yes/No (1/0)
Disease history	History of wheezing	Yes/No (1/0)
Family history of diseases	Allergy diseases	Yes/No (1/0)
	Asthma	Yes/No (1/0)
Assessment of atopic status	Protein	Yes/No (1/0)
	Fungus	Yes/No (1/0)
	Dust mite	Yes/No (1/0)
	Pollen or grass or tree	Yes/No (1/0)
	Total serum IgE values	Numerical
Physical examination	Whistling in the chest	Yes/No (1/0)
	Silent in the chest	Yes/No (1/0)
Inducements	Inducements	Yes/No (1/0)
Demographic information	Age (decimal in years)	Numerical

IgE, immunoglobulin E.

**Table 3** Statistics of missing features in different departments

Missing features	Pulmonology department (%)	Other departments (%)
History of wheezing	0.2	0.7
History of allergy	0	0.9
Family disease history	2.4	1.7
Allergen testing (for serum IgE)	31.81	65.23

IgE, immunoglobulin E.

feature value was set to 1. Features and their corresponding properties/values extracted from the EMR fields including cough, wheeze, tachypnea, chest tightness, history of wheezing, family history of asthma and allergy, ages, and various allergen test results, are listed in *Table 2*. The features extracted from the EMR fields were highly sparse with many empty entries. To reduce the level of sparsity, we limited the number of regular expressions. As a result, 16 regular expressions were developed for feature extraction from commonly used fields in real clinical settings (as shown in *Table 2*). The statistics of missing features for both datasets are also summarized in *Table 3*.

### Statistical analysis

In this study, four sets of statistical analysis between asthma positive and negative groups were analyzed using *t*-test within the Pulmonology Department, the non-Pulmonology Department, as well as across the Pulmonology and the non-Pulmonology Departments. As shown in *Tables 4* and *5*, most of the features extracted from the EHR data showed significant statistical difference ( $*P < 0.05$ ) between asthma positive and negative groups within both the Pulmonology Department and the non-Pulmonology Department, which confirmed that the features extracted were indeed

**Table 4** Statistical characteristics of subjects within different clinical departments

Characteristic	Pulmonology Department			Non-pulmonology Department		
	Positive group (N=1,624)	Negative group (N=2,137)	P	Positive group (N=337)	Negative group (N=1,786)	P
Chief complaints (mean ± SD)						
Cough	0.966±0.1825	0.907±0.2899	<0.0001*	0.938±0.2420	0.169±0.3749	<0.0001*
Wheezing	0.754±0.4305	0.358±0.4795	<0.0001*	0.707±0.4551	0.037±0.1882	<0.0001*
Tachypnea	0.064±0.2448	0.055±0.2284	0.2556	0.063±0.2437	0.011±0.1035	<0.0001*
Chest tightness	0.006±0.0782	0.001±0.0216	0.0014*	0.008±0.0880	0.004±0.0667	0.0198*
Disease history (mean ± SD)						
History of wheezing	0.770±0.4210	0.210±0.4074	<0.0001*	0.817±0.3867	0.035±0.1848	<0.0001*
Family history of diseases (mean ± SD)						
Allergy diseases	0.008±0.0891	0.002±0.0432	0.0055*	0.008±0.0880	0.001±0.0227	<0.0001*
Asthma	0.007±0.0856	0.003±0.0571	0.0780	0.013±0.1138	0.001±0.0292	<0.0001*
Assessment of atopic status (mean ± SD)						
Protein	0.108±0.3101	0.023±0.1497	<0.0001*	0.084±0.2774	0.019±0.1356	<0.0001*
Dust mite	0.055±0.2288	0.005±0.0682	<0.0001*	0.055±0.2279	0.003±0.0540	<0.0001*
Pollen or grass or tree	0.018±0.1324	0.003±0.0529	<0.0001*	0.024±0.1522	0.004±0.0627	<0.0001*
Total serum IgE	288.883±379.2370	79.546±301.1114	<0.0001*	166.857±427.6217	59.982±423.8480	<0.0001*
Physical examination (mean ± SD)						
Whistling in the chest	0.674±0.4687	0.370±0.4828	<0.0001*	0.688±0.4633	0.047±0.2120	<0.0001*
Demographic information (mean ± SD)						
Age	3.680±2.6891	2.176±2.8111	<0.0001*	3.605±2.7956	4.975±4.0442	<0.0001*

\*, P<0.05, with statistical difference between two groups. SD, standard deviation; IgE, immunoglobulin E.

discriminative. Among all asthma positive patients, the majority of the extracted features showed insignificant statistical difference (\*P>0.05) between the Pulmonology Department and the non-Pulmonology Department which confirmed that all asthma positive patients have similar characteristics. Among asthma negative patients, the majority of the extracted features showed significant statistical difference (\*P<0.05) between the Pulmonology Department and the non-Pulmonology Department, which also make sense since patients with diverse diseases should have different characteristics.

### Machine learning models

In this study, 80% of DataSet-1 from the Pulmonology Department which comprised 1,299 asthma cases and 1,709 negative cases was used as training data for modeling. The

remaining 20% comprising 325 positive and 428 negative cases was used as TestSet-1 for performance evaluation. Four machine learning based models, including CatBoost, Naïve Bayes, Support Vector Machines (SVM), and Logistic Regression (10-12) have been developed to identify asthma. CatBoost is an innovative ordered gradient boosting algorithm, which uses ordered target-based statistics for categorical features processing and permutation strategies to avoid prediction shift. Its base learner is an oblivious tree and each tree corresponds to a partition of the feature space. The model learns the feature space partition at each training iteration and finally obtains the aggregated data as a classification result. Naïve Bayes is a generative model that uses Bayes probabilities for calculating the classification total probabilities based on feature value combination; it assumes that each feature is independent. SVM aims at locating a hyper plane in the feature space with maximum

**Table 5** Statistical characteristics of subjects between asthma positive and negative groups

Characteristic	Positive groups			Negative groups		
	Pulmonology (N=1,624)	Non-pulmonology (N=337)	P	Pulmonology (N=2,137)	Non-pulmonology (N=1,786)	P
Chief complaints (mean ± SD)						
Cough	0.966±0.1825	0.938±0.2420	<0.0001*	0.907±0.2899	0.169±0.3749	<0.0001*
Wheezing	0.754±0.4305	0.707±0.4551	0.0005*	0.358±0.4795	0.037±0.1882	<0.0001*
Tachypnea	0.064±0.2448	0.063±0.2437	0.9313	0.055±0.2284	0.011±0.1035	<0.0001*
Chest tightness	0.006±0.0782	0.008±0.0880	0.5234	0.001±0.0216	0.004±0.0667	0.0060*
Disease history (mean ± SD)						
History of wheezing	0.770±0.4210	0.817±0.3867	0.0001*	0.210±0.4074	0.035±0.1848	<0.0001*
Family history of diseases (mean ± SD)						
Allergy diseases	0.008±0.0891	0.008±0.0880	0.9416	0.002±0.0432	0.001±0.0227	0.0324*
Asthma	0.007±0.0856	0.013±0.1138	0.0739	0.003±0.0571	0.001±0.0292	0.0034*
Assessment of atopic status (mean ± SD)						
Protein	0.108±0.3101	0.084±0.2774	0.0070*	0.023±0.1497	0.019±0.1356	0.1954
Dust mite	0.055±0.2288	0.055±0.2279	0.9481	0.005±0.0682	0.003±0.0540	0.1854
Pollen or grass or tree	0.018±0.1324	0.024±0.1522	0.1861	0.003±0.0529	0.004±0.0627	0.4279
Total serum IgE	288.883±379.2370	166.857±427.6217	<0.0001*	79.546±301.1114	59.982±423.8480	0.0413*
Physical examination (mean ± SD)						
Whistling in the chest	0.674±0.4687	0.688±0.4633	0.3321	0.370±0.4828	0.047±0.2120	<0.0001*
Demographic information (mean ± SD)						
Age	3.680±2.6891	3.605±2.7956	0.3694	2.176±2.8111	4.975±4.0442	<0.0001*

\*, P<0.05, with statistical difference between two groups. SD, standard deviation; IgE, immunoglobulin E.

margin. Logistic Regression aims to squeeze weighted sum combinations of feature values into categories. In our study, we applied a grid-search strategy with 5-fold cross validation on the training data to find the optimal hyper-parameters for all models including “maximum tree depth”, “maximum iteration number”, “learning rate”, “L2 leaf regularization”, and “early stopping rounds” for CatBoost; the “margin cost” parameter “C”, the kernel parameter “gamma”, and “kernel type” for SVM; and the “penalty strength” parameter “C” and “maximum iteration number” for Logistic Regression.

### Performance evaluation metrics

Algorithm performance was measured by two metrics, namely, the accuracy rate (Eq. [1]) and the area under the receiver operating curve (AUC). The accuracy rate

calculated the percentage of correctly predicted individuals among the whole test set, whereas the receiver operating curve (ROC) was generated by plotting the curve of sensitivity (Eq. [2]) *vs.* specificity (Eq. [3]).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad [1]$$

$$Sensitivity = \frac{TP}{TP + FN} \quad [2]$$

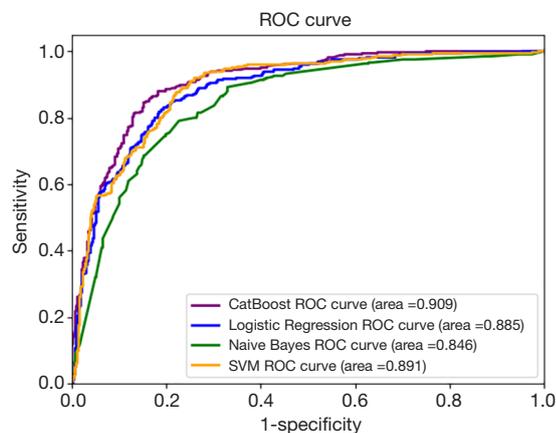
$$Specificity = \frac{TN}{TN + FP} \quad [3]$$

where TP, FP, TN, FN are true positive, false positive, true negative, and false negative rates, respectively. TP and TN represent correctly predicted positives and negatives with respect to the ground truth labels. FP and FN represent incorrectly predicted positives and negatives with respect to the ground truth labels.

**Table 6** Performance comparison on TestSet-1

Model	Accuracy (%)	AUC (%)
CatBoost	84.7	90.9
Logistic regression	80.5	88.5
SVM	80.3	89.1
Naïve Bayes	77.8	84.6

AUC, area under the receiver operating curve.

**Figure 1** Performance evaluation on TestSet-1. ROC, the receiver operating curve.

## Results

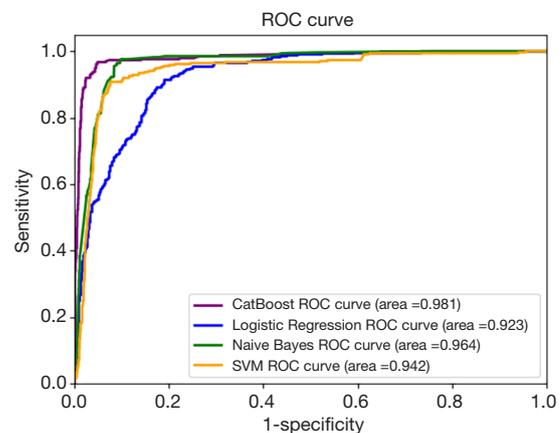
Two independent test sets were applied for performance evaluation. TestSet-1 consisted of 753 cases with 325 positive asthma cases and 428 negative cases from the Pulmonology Department. *Table 6* summarizes the overall performance of all four models on TestSet-1 where the CatBoost model outperformed all other models with an accuracy of 84.7% and an AUC of 90.0%. This result exceeded the second-best Logistic Regression and SVM by about 4% in accuracy and 2% in AUC. Naïve Bayes exhibited relatively lower scores, displaying less than 80% accuracy and an AUC of 84.62% respectively. *Figure 1* illustrates the AUC of these models more explicitly.

Similarly, *Table 7* and *Figure 2* demonstrate the overall performance of all four models on TestSet-2, composed of 2,123 cases (with 337 asthma-positive and 1,786 asthma-negative cases) from the non-pulmonology departments including the Cardiovascular Department, Endocrinology Department, Nephrology Department, Neurology Department, and Hematology Department. Since this

**Table 7** Performance comparison on TestSet-2

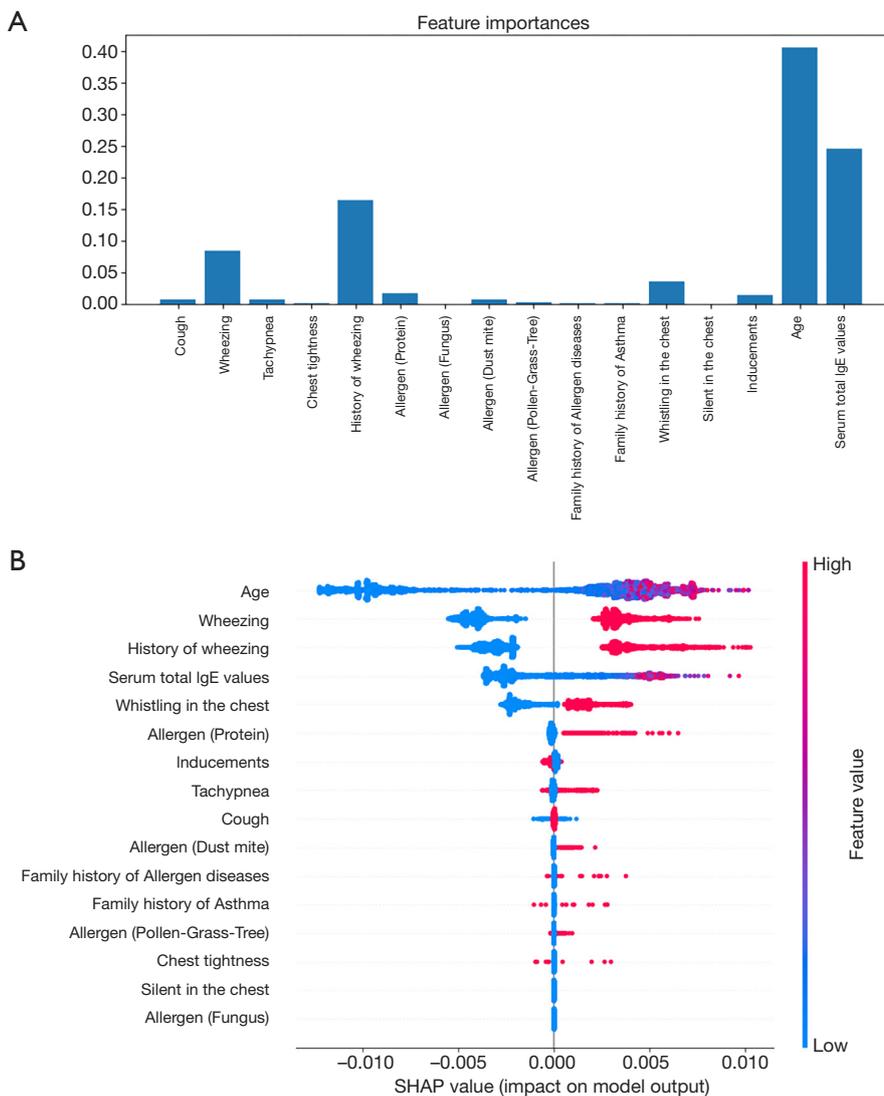
Model	Accuracy (%)	AUC (%)
CatBoost	96.7	98.1
SVM	91.8	94.2
Logistic Regression	88.6	92.3
Naïve Bayes	91.0	96.4

AUC, area under the receiver operating curve.

**Figure 2** Performance evaluation on TestSet-2. ROC, the receiver operating curve.

dataset was composed of data from other departments not involved in training, it was a perfect testbed to evaluate the robustness and generalization capacity of the model via quantitatively and qualitatively assessing its efficacy in a real clinical setting. As shown in *Table 7* and *Figure 2*, the CatBoost model displayed superior performance with an accuracy of 96.7% and an AUC of 98.1%. On TestSet-2, Naïve Bayes performed well with an accuracy of 91.0% and an AUC of 96.4%. The Logistic Regression was ranked last among these four models with an accuracy of 88.6% and an AUC of 92.3%.

In addition to accuracy and AUC, in this study, feature importance was also explored. This was calculated based on each feature's contribution to the model in performance. *Figure 3* shows the feature importance calculated by the CatBoost model. Age, considered seemingly irrelevant to asthma symptoms, was a highly correlated feature (*Figure 3A*). In clinical practice, pediatricians often consider wheezing and wheezing history with age. For example, a patient over the age of three exhibiting wheeze is more likely to have asthma than the patient's younger counterparts.



**Figure 3** Feature importance calculated by the CatBoost Model. SHAP, SHapley Additive exPlanation.

Wheezing and wheezing history was ranked as second and third most important features, exhibiting a combined weighted score of over 40%. While total serum IgE is an important feature in practice, it was not ranked among the top three most important features by CatBoost. This could be attributed to the fact that 40% of patients would not have been administered serum IgE tests, resulting in empty entries in the EMR. The degree of empty fields in this feature left a degree uncertainty, and potentially led to a lower ranking by the model. *Figure 3B* demonstrated polarity (positive or negative) impacts of features on model performance.

We also conducted ablation experiments on top feature bagging. As illustrated in *Table 8*, a superior performance

was observed with more features in general. However, the performance improvement was marginal on top 12 features and above, which is also demonstrated in the feature importance graph in *Figure 3*.

## Discussion

### Implications and findings

Our findings reveal that CatBoost clearly exhibited the highest accuracy and AUC compared to all other models on both test sets. CatBoost’s success may have be explained by its ability to process categorical features and modeling

**Table 8** Performance of top feature bagging

Top-N features	TestSet-1		TestSet-2	
	Accuracy (%)	AUC (%)	Accuracy (%)	AUC (%)
Top 4	84.1	90.3	95.0	96.9
Top 5	84.3	90.8	95.8	97.9
Top 8	84.5	90.8	96.6	98.0
Top 9	84.5	90.8	96.7	98.1
Top 12	85.0	90.8	96.3	98.1
All 16 features	84.7	90.9	96.7	98.1

AUC, area under the receiver operating curve.

feature combinations. Additionally, CatBoost's new capacity in undertaking feature combination increases its nonlinear modeling abilities. On TestSet-1, Naïve Bayes performed the worst, which was likely due to the feature independence assumption of this model. In practice, features from allergen test results are related to serum IgE, and thereby violate the independence assumption. On the other hand, Naïve Bayes performed well on TestSet-2; it is reasonable to conclude that as this model received higher negative label probability with more negative individuals, it had increased chances for negative predictions. In addition, the insensitivity of Logistic Regression to categorize features and its decreased ability to model nonlinear functions, along SVM's assumption of kernel space, might have contributed to their lower ranking compared to CatBoost on both data sets. Generally, all models demonstrated improved performance on TestSet-2. This was due to the presence of increased predictable negative instances of patients unlikely to have asthma, such as patients with no cough, wheeze, tachypnea, oppression in chest, history of wheeze, negative allergen test results, or family history of asthma and allergies.

### *Advantages of AI*

At present, the accuracy of Chinese primary pediatricians in diagnosing asthma is relatively low. One study of pediatricians in Jinhua City, Zhejiang Province, reported the average accuracy of asthma diagnosis among pediatricians at all professional levels to be only 59.29%. In addition, the diagnostic accuracy of primary units at lower professional levels (e.g., level 1 and level 2) was lower than that of level

3 (4). Therefore, if this AI system is routinely applied to the work of primary pediatricians, it will greatly increase the accuracy of their asthma diagnosis, and thus, significantly improve the management of asthma in Chinese children.

### *Limitations*

Currently, this system is only suitable for identifying the existence of asthma, and it cannot provide disease severity classification or treatment suggestions. Furthermore, it is also unable to prompt clinicians to complete medical records and auxiliary examinations.

### *Future expectations*

The complete AI system including the NLP-based feature extraction module and the machine learning-based asthma identification module can be integrated into hospitals' information and data management systems as a whole to identify asthma in pediatric inpatient setting, through extracting data from EMR as inputs for the developed AI system. With a questionnaire-based interactive interface for feature inputs, it could allow the AI model to be applied as an efficient screening tool in the outpatient service system to identify more patients with a mild level asthma.

At present, the developed asthma identification model has been integrated into a smart phone APP with a questionnaire-based interactive interface for feature inputs and applied in the outpatient service in the Children's Hospital, Zhejiang University School of Medicine, China.

On the basis of a pre-diagnosis function, systems focused on auxiliary treatment and follow-up could be simultaneously developed to utilize AI-assisted methods as complete management tools for asthma patients.

### **Conclusions**

The AI model can quickly and accurately identify children with asthma, which can aid primary pediatricians in making more precise diagnoses, whilst further preventing undetected cases. Our findings are of great clinical value and practical significance in improving the asthma control level of children in China, optimizing medical resources, and decreasing the abuse of antibiotics and systemic glucocorticoids.

## Acknowledgments

*Funding:* This study was supported in part by grants from the National Natural Science Foundation of China (No. 62076218), the Zhejiang Province Research Project of Public Welfare Technology Application (LGF18H260004), and the National Natural Science Foundation of China (No. 81870023 and 81000765).

## Footnote

*Data Sharing Statement:* Available at <http://dx.doi.org/10.21037/atm-20-2501a>

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/atm-20-2501a>). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Study approval was granted by The Institutional Review Board (IRB) of the Medical Ethics Committee of Children's Hospital, Zhejiang University School of Medicine, China (IRB Approval ID: 2020-IRB-039). The procedures were performed in accordance with the Declaration of Helsinki (as revised in 2013) and relevant guidelines and regulations. Informed consent was obtained after the procedure was fully explained to all participants and their legal guardians.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. National Cooperative Group on Childhood Asthma;

- Institute of Environmental Health and Related Product Safety, Chinese Center for Disease Control and Prevention; Chinese Center for Disease Control and Prevention. Third nationwide survey of childhood asthma in urban areas of China. *Zhonghua Er Ke Za Zhi* 2013;51:729-35.
2. Sha L, Liu C, Shao M, et al. Ten years comparison of diagnosis and treatment of asthma in urban children in China. *Zhonghua Er Ke Za Zhi* 2016;54:182-6.
  3. Wong GW, Kwon N, Hong JG, et al. Pediatric asthma control in Asia: phase 2 of the Asthma Insights and Reality in Asia-Pacific (AIRIAP 2) survey. *Allergy* 2013;68:524-30.
  4. McGeachie MJ, Yates KP, Zhou X, et al. Patterns of Growth and Decline in Lung Function in Persistent Childhood Asthma. *N Engl J Med* 2016;374:1842-52.
  5. Ye X, Xu C, Wang Z. Cognition of asthma guidelines among doctors in Jinhua city. *Chinese J of Pub Health* 2015;31:804-6.
  6. Howard R, Rattray M, Prosperi M, et al. Distinguishing Asthma Phenotypes Using Machine Learning Approaches. *Curr Allergy Asthma Rep* 2015;15:38.
  7. Prosperi MC, Marinho S, Simpson A, et al. Predicting phenotypes of asthma and eczema with machine learning. *BMC Med Genomics* 2014;7 Suppl 1:S7.
  8. Tomita K, Nagao R, Touge H, et al. Deep learning facilitates the diagnosis of adult asthma. *Allergol Int* 2019;68:456-61.
  9. Pediatric Branch of Chinese Medical Association, the guide of Children's bronchial asthma diagnosis and prevention. *China J Pediatr* 2016;54:3.
  10. Prokhorenkova L, Gusev G, Vorobev A, et al. CatBoost: unbiased boosting with categorical features. *Adv in Neural Infor Proc Sys*, 2018.
  11. Cortes C, Vapnik V. Support-vector networks. *Machine Learning* 1995;20:273-97.
  12. Hand D, Yu K. Idiot's Bayes – not so stupid after all? *International Statistical Review* 2001;69:385-99.

(English Language Editors: E. Tan and J. Gray)

**Cite this article as:** Yu G, Li Z, Li S, Liu J, Sun M, Liu X, Sun F, Zheng J, Li Y, Yu Y, Shu Q, Wang Y. The role of artificial intelligence in identifying asthma in pediatric inpatient setting. *Ann Transl Med* 2020;8(21):1367. doi: 10.21037/atm-20-2501a