Application and Prospects of Artificial Intelligence Technology in Early Screening of Chronic Obstructive Pulmonary Disease at Primary Healthcare Institutions in China

Xu Yang
Department of General Practice, Donghuashi Community Health Service Center, Beijing, People’s Republic of China

Correspondence: Xu Yang, Department of General Practice, Donghuashi Community Health Service Center 33 Donghuashi Street, Dongcheng District, Beijing, 100062, People’s Republic of China, Tel +86 13693088802, Fax +86 67120077, Email xuyang8799@163.com

Abstract: Chronic Obstructive Pulmonary Disease (COPD), as one of the major global health threat diseases, particularly in China, presents a high prevalence and mortality rate. Early diagnosis is crucial for controlling disease progression and improving patient prognosis. However, due to the lack of significant early symptoms, the awareness and diagnosis rates of COPD remain low. Against this background, primary healthcare institutions play a key role in identifying high-risk groups and early diagnosis. With the development of Artificial Intelligence (AI) technology, its potential in enhancing the efficiency and accuracy of COPD screening is evident. This paper discusses the characteristics of high-risk groups for COPD, current screening methods, and the application of AI technology in various aspects of screening. It also highlights challenges in AI application, such as data privacy, algorithm accuracy, and interpretability. Suggestions for improvement, such as enhancing AI technology dissemination, improving data quality, promoting interdisciplinary cooperation, and strengthening policy and financial support, aim to further enhance the effectiveness and prospects of AI technology in COPD screening at primary healthcare institutions in China.

Keywords: chronic obstructive pulmonary disease, primary healthcare institutions, artificial intelligence, high-risk group screening, data privacy

Introduction
Chronic Obstructive Pulmonary Disease (COPD) is a heterogeneous lung disease that causes persistent and progressive airflow obstruction, and it has become one of the major threats to global health. Chronic respiratory symptoms caused by COPD, such as difficulty breathing, coughing, and expectoration, not only severely affect the quality of life of patients but also pose an urgent public health issue due to its high mortality rate and economic burden.1 It is estimated that approximately 6% of deaths worldwide are attributed to COPD,2 and the prevalence and mortality rates in China are particularly significant. With the aging population and continuous exposure to risk factors of COPD, the prevalence of COPD in Chinese adults over 40 years old has sharply risen to 13.7%3 in recent years, affecting nearly 100 million adults.4

COPD is a preventable and treatable disease, and early diagnosis is crucial for controlling disease progression and improving prognosis. Screening and early intervention in high-risk groups can delay disease progression and improve outcomes. However, the current awareness and diagnosis rates of COPD remain low. The survey results of the “China Adult Lung Health Study” show that only 10% of the study subjects were aware of the disease name COPD, and less than 10% had undergone lung function tests.4 Primary healthcare institutions play an important role in providing basic healthcare to residents, particularly in the discovery and early diagnosis of COPD among high-risk groups.5 With the...
development and application of Artificial Intelligence (AI) technology, these institutions are expected to enhance the efficiency and accuracy of COPD screening.

The application of artificial intelligence in medicine, especially within the field of respiratory systems, is gradually becoming a focal point. In this domain, particularly in the diagnosis and management of chronic obstructive pulmonary disease (COPD), the development of digital health technologies and big data science has brought unprecedented opportunities to the medical domain. Artificial intelligence has not only demonstrated high accuracy in interpreting pulmonary function test results, but also shown immense potential in medical imaging. Radiomics techniques extract a plethora of features from medical images using machine learning methods, providing new avenues for the early diagnosis and risk assessment of COPD. These techniques are capable of detecting early-stage patients that are difficult for the human eye to identify. Furthermore, deep learning techniques, particularly deep convolutional neural networks (CNNs), have exhibited outstanding performance in automatically interpreting lung CT images and diagnosing obstructive lung diseases, offering powerful tools for the diagnosis and management of COPD. Additionally, with the continuous evolution of digital health technologies and the development of big data science, an increasing number of studies are focusing on leveraging machine learning and radiomics techniques, combined with large-scale clinical data, to improve early screening, diagnosis, and treatment plans for COPD. The aim is to achieve personalized medicine and enhance patient quality of life. In summary, the development of digital health technologies and big data science brings new hope for the diagnosis and management of COPD, while the application of artificial intelligence provides strong support for achieving personalized medicine and improving medical efficiency. Exploring the application of AI in enhancing disease screening and early diagnosis capabilities in primary healthcare institutions is of key significance in addressing this major public health challenge.

Methods
A comprehensive search of Chinese and English databases was conducted, including the CNKI and Wanfang databases within China, as well as the Embase and PubMed databases, to retrieve relevant data. Additionally, potential studies meeting the criteria were identified through manual searches. The research scope was limited to Chinese and English literature published between January 1, 2019, and January 1, 2024. Keywords used included both Chinese and English terms such as Chronic Obstructive Pulmonary Disease (COPD), Early Screening, COPD Screening, Artificial Intelligence, Primary Healthcare Institutions, Healthcare Data Analysis, and Healthcare in China. Following database searches, articles were manually screened to select those that met the criteria for inclusion in this study. Specifically, 12 articles were selected from the CNKI database, 4 articles from the Wanfang database, 19 articles from the English PubMed database, and 15 articles from the Embase database.

Target Population for Screening Strategies
The high-risk group for COPD includes individuals with: a long history of smoking (over 20 pack-years); recurrent chest infections; early life events, such as frequent respiratory infections in earlier years; increasing age, especially those over 40; and household air pollution. In low and middle-income countries, nearly 3 billion people use biomass and coal as the main sources of energy for cooking, heating, and other household needs, thus posing a substantial global risk to a large population.

Current COPD Screening Methods and Tools
Symptom Assessment and Physical Examination
This involves taking a medical history, focusing on chronic cough, expectoration, difficulty breathing, chest tightness, and other symptoms. Additionally, assessing the patient’s quality of life is important to understand the impact of symptoms on daily activities. This assessment can predict the occurrence and progression of COPD. Some primary healthcare institutions can screen high-risk groups for COPD based on symptom and quality of life assessments. Physical examination is crucial in the care of COPD patients, but in terms of diagnosis, its sensitivity and specificity are
relatively low due to the fact that signs of airflow obstruction typically do not appear until lung function is significantly impaired.¹

**Pulmonary Function Tests**

These include spirometry and peak flow meter tests. Spirometry is the gold standard for diagnosing COPD.¹ The test measures the volume of air expelled by the patient in a certain amount of time, as well as the total volume of air that can be exhaled. Spirometry can assess airflow limitation, characterized by a reduced ratio of forced expiratory volume in one second (FEV1) to forced vital capacity (FVC).¹

**PEF Measurement**

Peak Expiratory Flow (PEF) measurement assesses the speed of forced exhalation, representing the peak flow rate of air during exhalation. It is used as a screening tool to identify airflow limitation. Peak flow meter tests measure the maximum speed of airflow expelled from the lungs, providing quick information about changes in lung function, but they are less diagnostic compared to spirometry.¹²

**Imaging Studies**

Chest X-rays, while not decisive for the diagnosis of COPD, can help rule out other lung diseases such as tuberculosis or lung cancer. High-Resolution Computed Tomography (HRCT) is used in some cases for more detailed assessment of lung structure, especially in patients with atypical X-ray findings.¹³

**Biomarkers and Blood Tests**

Although there are no specific biomarkers for the screening of COPD at present, certain blood tests (such as blood gas analysis, C-reactive protein) can provide information about the patient’s inflammatory status and oxygenation.¹⁴

**Questionnaire Surveys**

Questionnaires are an economical and convenient method for early screening of COPD. The GOLD guidelines recommend that primary healthcare institutions use questionnaires for active case finding. Questionnaire surveys are significant for primary healthcare institutions in identifying high-risk groups for COPD. Various COPD screening questionnaires have been clinically developed, such as the CAPTURE questionnaire,¹⁵ Chronic Obstructive Pulmonary Disease Screening Questionnaire (COPD-SQ),¹⁶ etc. The COPD-SQ was developed by a research team using data collected in 2002 from 19,800 Chinese subjects aged 40, including 7 items: age, smoking pack-years, body mass index, cough, difficulty breathing, family history of respiratory diseases, and exposure to biomass smoke from cooking. The scale has a COPD diagnosis sensitivity of 60.6%, specificity of 85.2%, and an accuracy rate of 82.7%.¹⁷ The widely used screening questionnaires for COPD are often based on studies conducted on populations from other countries, and whether they are applicable to China remains to be further explored. In 2023, a small-scale COPD screening questionnaire was applied in community-based primary healthcare institutions in China. The results showed that compared to diagnostic pulmonary function tests, the COPD-SQ scale demonstrated a specificity of 59.69%, accuracy (approximate index of 0.246), and an area under the curve of 0.744 (95% CI 0.650, 0.837).¹⁸

**Application of AI in COPD Screening**

Diagnostic Assistance

AI-Assisted Pulmonary Function Diagnosis: Pulmonary function tests play a crucial role in diagnosing COPD. Lung parameters measured by these tests, such as Forced Vital Capacity (FVC) and Forced Expiratory Volume in One Second (FEV1), are the gold standards for confirming COPD. Besides, pulmonary function is also used for staging COPD patients, differentiating other respiratory diseases with similar symptoms like asthma in early diagnosis, and monitoring disease progression in long-term management. Regular pulmonary function tests help in monitoring changes in the condition, assessing future health risks, evaluating treatment effectiveness, and timely adjusting treatment plans.
Therefore, standardized and accurate pulmonary function test results are vital for comprehensive COPD management. Artificial Intelligence can assist in interpreting results and diagnosing pulmonary function tests. Marko Topalovic and others compared the accuracy of AI software and pulmonologists in interpreting pulmonary functions. The results showed that AI software had a match rate of 100% in interpreting test results and a diagnostic accuracy rate of 82%. In contrast, pulmonologists had a match rate of 74.4% and a diagnostic accuracy rate of 44.6%. This study highlights the potential of AI in improving the accuracy of pulmonary function interpretation and its possibility as a decision-support tool in clinical practice.

AI Technology in Recognizing Imaging for Early Diagnosis: AI technology has been used to analyze lung CT scans and X-ray images to identify early images of COPD. Through deep learning algorithms, it can accurately pinpoint changes in lung structure, assisting doctors in early and precise detection of COPD. The process of imaging omics primarily involves importing imaging data, utilizing software such as ITK-SNAP and 3D Slicer to delineate regions of interest, extracting imaging omics features, and establishing machine learning models based on the selected imaging omics features, such as logistic regression (LR) model, support vector machine (SVM) model, K-nearest neighbor (KNN) algorithm, etc. Finally, the performance of the machine learning model is evaluated. For instance, in 2022, Li Z and others developed a Graph Convolutional Network (GCN) for early detection of COPD. This method utilized chest computed tomography image data from the publicly available Danish Lung Cancer Screening Trial database. The GCN model achieved an accuracy of 0.77 and an area under the curve of 0.81.

AI Technology in Integrating and Analyzing Data for Early Diagnosis
AI can efficiently process and analyze large volumes of electronic health records, mine data, and extract key information to help identify risk factors for COPD and predict high-risk groups, such as those with a history of smoking or occupational exposure. By analyzing patients’ past medical records and vital signs, AI can build predictive models to forecast the development trends of COPD, further assisting doctors in making better personalized screening decisions. For example, Clinical Decision Support Systems (CDSS) are defined as electronic systems that directly assist in clinical decision-making, which can help generate specific evaluations or recommendations for patients and present them to clinical doctors for reference. Although the World Health Organization (WHO) has identified the development of health information systems and digital technologies (including CDSS) as one of the priorities for strengthening primary healthcare, the application of CDSS for evaluating patients with respiratory distress in primary healthcare and outpatient services in China has not yet become widespread. Another study published in “CHEST” explored the use of AI as a predictive tool to identify high-risk COPD patients. Certain AI systems use sensor data collected from wearable devices to assist in diagnosing COPD by analyzing changes in breathing patterns, such as frequency, depth, and rhythm.

Challenges and Limitations of AI in COPD Screening
While Artificial Intelligence (AI) offers many potential advantages in the screening of Chronic Obstructive Pulmonary Disease (COPD), there are also significant challenges and limitations.

Data Privacy: In terms of protecting sensitive information, safeguarding personal privacy is crucial when handling patients’ medical records and diagnostic data. It is essential to ensure that data collection and processing comply with relevant privacy laws and standards, such as the European Union’s General Data Protection Regulation (GDPR). Regarding data security, strong safety measures must be implemented to prevent data breaches and unauthorized access. This includes data encryption, secure data storage and transmission, and strict control over data access. In China, the privacy of medical information data is covered by the Personal Information Protection Law implemented in 2021 and the Data Security Law, both of which involve the secure processing of data, including medical data. These laws regulate data processing activities to ensure data security and prevent data leakage. Furthermore, the “National Health and Medical Big Data Standards, Security, and Service Management Measures (Trial)”, specifically aimed at data security management in the healthcare industry, covers regulations on the collection, storage, transmission, sharing, and destruction of medical data. These laws and regulations constitute the legal framework for the protection of medical information data privacy in China, requiring medical institutions and related companies to strictly adhere to the relevant
provisions of privacy protection and data security. As technology develops and the application of data expands, these regulations may be continuously updated and improved to meet new challenges and needs.

Algorithm Accuracy: Data Quality and Representativeness: The accuracy of AI models highly depends on the quality and representativeness of the data used for training. In a study on the diagnostic value of artificial intelligence assisted diagnosis of COPD in China, the sensitivity of the artificial intelligence robot questionnaire for screening COPD was 76.11%, the specificity was 84.76%, the Jordan index was 60.87%, and the area under the ROC curve was 0.858. Based on CT imaging, COPD recognition was performed using a multi instance learning logic classifier on a dataset of 100 COPD patients and 100 healthy subjects, resulting in an AUC of 0.742. If the data is biased or incomplete, it can lead to decreased performance of the model in real-world applications. Different populations (such as different ethnicities, age groups, or geographical locations) may exhibit different phenotypic variations of COPD. AI models need to be able to adapt to these differences. For example, in a study, six machine learning classifiers, LR, SVM, KNN, RandomForest, ExtraTrees, and XGBoost, were trained to construct an imaging omics model. The performance of the imaging omics models constructed based on these six machine learning models varies, with XGBoost having the best diagnostic value for COPD. However, the sample size used for the establishment of imaging omics models in China is relatively small, and further expansion of the sample size is needed for validation. Moreover, most studies are single center studies, and external independent tests need to be conducted in multiple centers. A 2022 study suggested that creating a specific extension for quality assessment tools for AI research in diagnostic accuracy could help safely translate AI tools into clinical practice.

Interpretability Issues: Many AI models, especially deep learning models, are considered “black boxes” because their decision-making processes are difficult to explain. This is particularly important in the medical field, where both doctors and patients may need to understand the basis of the model’s decisions. When using AI in medical decision-making, it is necessary to clarify issues of responsibility attribution and compliance. For instance, if an AI model’s diagnostic or treatment recommendation is incorrect, it should be clear where the responsibility lies.

**Conclusion and Outlook**
The application of AI in the screening of COPD at primary healthcare institutions in China is gradually unfolding. AI technology has shown potential in disease screening, condition monitoring, data analysis, and patient education, particularly in handling large volumes of electronic health records, data mining, and extracting key information. AI provides an effective tool for identifying COPD risk factors and predicting high-risk groups. Additionally, AI has significant advantages in enhancing telemedicine services, designing personalized medical plans, and improving the efficiency of public health strategies.

To effectively enhance the application of AI technology in COPD screening at primary healthcare institutions in China, training and education can be implemented to improve medical personnel’s understanding and ability to use AI, ensuring its effective application in COPD screening and management. Secondly, improving the quality of data collection and processing, including ensuring the accuracy and comprehensiveness of collected health data and strengthening data privacy and security, is essential so that AI can be more effectively used for data analysis and decision support. Furthermore, promoting interdisciplinary collaboration between the healthcare sector and fields such as computer science and data analysis is crucial for developing and implementing effective AI solutions. Lastly, the government and related institutions need to strengthen policy support and financial investment to promote the widespread use and application of AI technology in primary healthcare institutions. By implementing these comprehensive measures, the application of AI in COPD screening at primary healthcare institutions in China will be significantly enhanced, better addressing this major public health challenge.

**Acknowledgments**
Funded by Capital’s Funds for Health Improvement and Research (CFH 2022-4-7014).

**Disclosure**
The author reports no conflicts of interest in this work.
References


