



# Automated valvular heart disease detection using heart sound with a deep learning algorithm

Zihan Jiang<sup>a,1,3</sup>, Wenhua Song<sup>b,1,3</sup>, Yonghong Yan<sup>c,3</sup>, Ao Li<sup>c,3</sup>, Yujing Shen<sup>a,3</sup>, Shouda Lu<sup>c,3</sup>, Tonglian Lv<sup>b,3</sup>, Xinmu Li<sup>b,3</sup>, Ta Li<sup>c,3</sup>, Xueshuai Zhang<sup>c,3</sup>, Xun Wang<sup>d,3</sup>, Yingjie Qi<sup>a,3</sup>, Wei Hua<sup>a,3</sup>, Min Tang<sup>a,2,\*</sup>, Tong Liu<sup>b,2,\*</sup>

<sup>a</sup> Arrhythmia Center, State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100037, China

<sup>b</sup> Tianjin Key Laboratory of Ionic-Molecular Function of Cardiovascular Disease, Department of Cardiology, Tianjin Institute of Cardiology, Second Hospital of Tianjin Medical University, Tianjin 300211, China

<sup>c</sup> Key Laboratory of Speech Acoustics and Content Understanding, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China

<sup>d</sup> Kaiserslautern Intelligent Manufacturing School, Shanghai Dianji University, Shanghai 200240, China

## ARTICLE INFO

### Keywords:

Machine learning  
Neural networks  
Valvular heart disease  
Heart sound  
Physical examination

## ABSTRACT

**Background:** Insufficient clinicians' auscultation ability delays the diagnosis and treatment of valvular heart disease (VHD); artificial intelligence provides a solution to compensate for the insufficiency in auscultation ability by distinguishing between heart murmurs and normal heart sounds. However, whether artificial intelligence can automatically diagnose VHD remains unknown. Our objective was to use deep learning to process and compare raw heart sound data to identify patients with VHD requiring intervention.

**Methods:** Heart sounds from patients with VHD and healthy controls were collected using an electronic stethoscope. Echocardiographic findings were used as the gold standard for this study. According to the chronological order of enrollment, the early-enrolled samples were used to train the deep learning model, and the late-enrollment samples were used to validate the results.

**Results:** The final study population comprised 499 patients (354 in the algorithm training group and 145 in the result validation group). The sensitivity, specificity, and accuracy of the deep-learning model for identifying various VHDs ranged from 71.4 to 100.0%, 83.5–100.0%, and 84.1–100.0%, respectively; the best diagnostic performance was observed for mitral stenosis, with a sensitivity of 100.0% (31.0–100.0%), a specificity of 100% (96.7–100.0%), and an accuracy of 100% (97.5–100.0%).

**Conclusions:** Based on raw heart sound data, the deep learning model effectively identifies patients with various types of VHD who require intervention and assists in the screening, diagnosis, and follow-up of VHD.

## 1. Introduction

Valvular heart disease (VHD) is a group of diseases with high morbidity and mortality that critically affects patients' quality of life. The incidence of VHD is becoming more prevalent with increasing life expectancy and is approximately 13.3 % in older patients.[1]

Simultaneously, annual VHD death rates are increasing, and deaths due to VHD are expected to double in 25 years.[2].

Most patients with VHD are only diagnosed after developing complications, such as heart failure, which has a poor disease prognosis because VHD has a protracted asymptomatic period.[3] Additionally, the cost of medical care for patients increases with the severity of VHD;

\* Corresponding authors at: Tianjin Key Laboratory of Ionic-Molecular Function of Cardiovascular Disease, Department of Cardiology, Tianjin Institute of Cardiology, Second Hospital of Tianjin Medical University, No. 23, Pingjiang Road, Hexi District, Tianjin 300211, China (T. Liu). Arrhythmia Center, State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, No.167 North Lishi Road, Xicheng District, Beijing 100037, China (M. Tang).

E-mail addresses: [doctortangmin@yeah.net](mailto:doctortangmin@yeah.net) (M. Tang), [liutongdoc@126.com](mailto:liutongdoc@126.com) (T. Liu).

<sup>1</sup> These authors have contributed equally to this work and share first authorship.

<sup>2</sup> These authors have contributed equally to this work and share corresponding authorship.

<sup>3</sup> This author takes responsibility for all aspects of the reliability and freedom from bias of the data presented and their discussed interpretation.

<https://doi.org/10.1016/j.ijcha.2024.101368>

Received 23 October 2023; Received in revised form 8 February 2024; Accepted 17 February 2024

2352-9067/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

[4] therefore, a timely diagnosis is key to improving patient outcomes and reducing healthcare costs.

Delays in VHD diagnosis and treatment can be caused by poor auscultation skills and underuse of auscultation by the clinician. [5,6] It has been discovered that medical professionals, including medical students, lack experience in auscultation, and the sensitivity and specificity of identifying heart murmurs are only 35–69 %. [6,7] Echocardiography is the primary method to confirm VHD and assess the disease severity. However, the longer examination time, dependence on specialist physicians for interpretation, and subjective factors, such as physician experience, make it unsuitable for routine screening and assessment. Therefore, it is vital to develop objective tests for VHD diagnosis that are easy to perform, reproducible, and accessible.

Artificial intelligence (AI)-assisted auscultation technology has offered a feasible solution for physicians' inexperience in recent years. Previous studies have shown that AI can distinguish normal heart sounds from heart murmurs with a sensitivity of 78.5–88.5 % and a specificity of 81.5–98.33 %. [8–10] The aforementioned dichotomous heart sound examinations cannot diagnose VHD because they can only distinguish between normal and abnormal heart sounds. Additionally, AI can determine the phonocardiograms of patients with aortic, mitral, and mitral regurgitation; [11] however, no phonocardiograms from patients with mixed valve diseases were included. Notably, most studies used processed heart sound signals or phonocardiogram signals rather than raw heart sound data collected in clinical settings; therefore, they did not reflect the actual clinical application of such techniques.

We analyzed and compared the heart sounds of patients with VHD of different severities and those of healthy individuals using raw heart sound signals collected in clinical settings. Additionally, patients with VHD requiring intervention were identified via algorithms to achieve an automatic diagnosis of VHD and to assess the disease severity for the diagnosis of VHD.

## 2. Methods

### 2.1. Study design

This study used a multisite diagnostic test. The participating institutions were Fuwai Hospital, National Center for Cardiovascular Diseases (Beijing, China), the Department of Cardiology, Tianjin Institute of Cardiology, Second Hospital of Tianjin Medical University (Tianjin, China), and the Institute of Acoustics, Chinese Academy of Sciences (Beijing, China). Both medical institutions participating in the study were tertiary care hospitals.

This is a prospective study. Patients with VHD were prospectively enrolled in each study sites. Echocardiographic findings were used as the gold standard for this study. Patients with VHD were defined as those who underwent echocardiography to confirm VHD diagnosis; the American Society of Echocardiography guidelines were used as the diagnostic criteria. [12] Healthy controls were outpatients or inpatients diagnosed without valve disease, congenital heart disease, or ventricular wall tumors by electrocardiography, echocardiography, or healthy volunteers who volunteered to participate in the study. According to VHD management guidelines, [13] samples were separated into "intervention required" and "no intervention required" groups based on surgical or percutaneous valve replacement indications. Heart sound collection was done by specialist nurses, all of whom had undergone uniform training in heart sound collection before the start of the study.

This study was divided into two phases: the AI algorithm model-building and validation periods. The timing of sample enrollment was used as the basis for grouping. A VHD heart sound diagnosis system was constructed and debugged using participant data from the model building. We noise-reduced the raw heart sound data, extracted features from the denoised audio, and trained the model using the features. After debugging the AI heart sound diagnostic system, the sensitivity and specificity of the validation group were tested (Fig. 1).

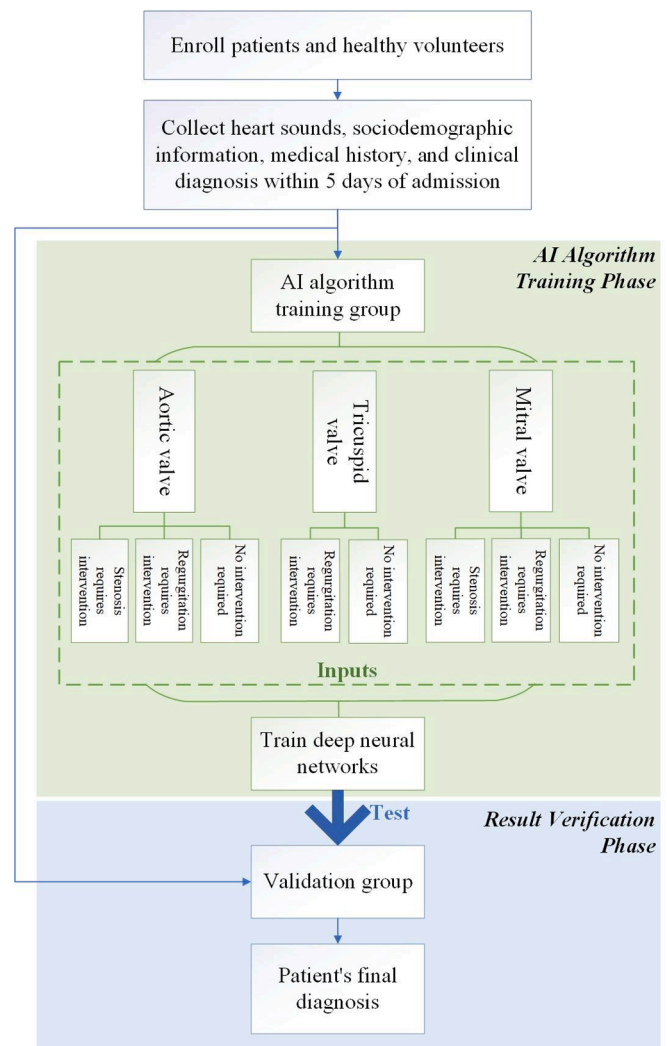


Fig. 1. Study design.

Regarding sample size estimation, previous studies reported algorithm sensitivities of 93–95 % [14,15] for detecting normal and abnormal heart sounds. Therefore, we assumed a sensitivity and specificity of 0.9 for the detection of clinically significant VHD and a sample size of 35 participants in the study and control groups ( $\alpha = 0.05$ , estimation error = 0.1).

The protocol was conducted in accordance with the Declaration of Helsinki and approved as a minimal-risk study by the institutional review boards of the participating sites (approval numbers: KY2021K100 and 2020-ZX14). Informed consent was obtained from all patients.

### 2.2. Participant enrollment

- (1). The following participants were included:
  - (i). Patients aged 18–80 years;
  - (ii). Patients with VHD, defined as those with at least one of the following diseases diagnosed using echocardiography: aortic stenosis, aortic valve insufficiency, mitral stenosis, mitral valve insufficiency, and tricuspid valve insufficiency;
  - (iii). Healthy controls, defined as participants with a confirmed absence of VHD and structural heart disease using echocardiography;
  - (iv). Patients who consented to participate in this study and signed an informed consent form.
- (2). Exclusion criteria were as follows:

- (i). Congenital heart disease and structural abnormalities in the heart that may produce disturbances in heart sounds, including atrial and ventricular septal defects and ventricular aneurysms;
- (ii). Pacemaker recipients;
- (iii). Unstable disease conditions that did not allow the patient to remain supine;
- (iv). Known allergic reactions to medical silicone rubber products;
- (v). Pregnancy.

### 2.3. Data collection

Sociodemographic information and the medical history were collected as baseline data. Echocardiographic and electrocardiographic data of all participants were extracted from the medical records.

Heart sound data was collected by specialized nurses who underwent a standardized training in the acquisition of heart sounds prior to the start of the current study. The specialized nurses collected heart sounds within 5 days of admission using an electronic stethoscope based on the Android system, which has been reported previously. [16] Patients were asked to remain quiet during data collection, and no talking or eating was permitted. Heart sounds were collected from the aortic valve auscultation area, mitral valve auscultation area, tricuspid valve auscultation area, and Erb's point for 30–120 s. The collected heart sounds were saved in a 16-bit single-channel wav format, with a 4,000-Hz sampling frequency, and were uploaded to the cloud platform for subsequent analysis. A study on the design and use of the electronic stethoscope and heart sound storage cloud platform has been reported. [16].

### 2.4. Heart sound signal processing and deep learning model construction

Automated noise reduction and feature extraction methodologies for heart sounds have been reported. [17] In the first step, the raw heart sound data were fed into a noise reduction program; an optimal modified log-spectral amplitude estimator combined with wavelet noise reduction was used. Notably, wavelet noise reduction operates by a threshold noise reduction technique. [18] The processed coefficients are then inversely transformed to reconstruct the noise-reduced signal.

In the second step, feature extraction is performed on the audio after heartbeat noise reduction is completed. Next, each sample was sliced into small 4-s segments, which could include at least two or more cardiac cycles. The window length of the frame was 25 ms, and the frameshift was 15 ms. Next, we employed the Mel-frequency cepstral coefficients (MFCCs) and root-mean-square energy for the features. First, the object signal was pre-emphasized, framed, and windowed and a fast Fourier transform was applied. Subsequently, the power spectrum was calculated, and the obtained power spectrum was passed through a triangular band-pass filter. The filtered output was converted into a logarithmic form using the Mel-domain versus linear frequency relationship. Finally, a discrete cosine transform was performed to obtain the MFCCs. The energy calculation in the time domain was based on the amplitude. The amplitude of a speech segment was used to deduce the root mean square, which was the result of squaring the sum of squares of N terms divided by N.

In the third step, the extracted features were simultaneously fed into the deep learning model (Supplementary Figure S1). We established three algorithm models for the aortic, mitral, and tricuspid valves. Each model eventually output heart sounds corresponding to the valve region as stenosis, regurgitation, or no intervention. For each algorithm, the softmax function was used to analyze the output layer of the model. The training process used the Adam optimizer with an initial learning rate of 0.001, and the learning rate decreased by 0.01 with every 10 epochs until the final model converged. To prevent overfitting, we used an early stopping strategy, and the model stopped training when its validation

loss stopped decreasing for 5 consecutive epochs. The detailed information about the architecture of deep learning was summarized in Supplementary Table S1a to Table S1c. A brief introduction of “Soft-max” and “Adam” was attached in the Supplementary.

The automation of the process is a key feature, as there is no requirement for manual intervention or correction during the signal processing, denoising, and deep learning model feeding stages. Additionally, the preprocessing steps have demonstrated a certain level of universality, making them applicable to a wide range of datasets and scenarios.

### 2.5. Statistical analysis

For continuous data, the mean  $\pm$  standard deviation was used for normal distributions and the median (quartiles) for skewed distributions. The percentage composition ratios represent categorical variables. The *t*-test, Kruskal–Wallis nonparametric test, and chi-square test were used to compare the groups. The sensitivity, specificity, and accuracy were used to describe the diagnostic efficacy of the AI system.

Multiple logistic regression (forward stepwise: likelihood ratio) was used to identify factors affecting diagnostic consistency. The 95 % confidence intervals (CI) for the proportions were calculated according to the efficient score-based method (corrected for continuity) described by Newcombe. [19]  $P < 0.05$  was considered statistically significant. Statistical analyses were performed using SPSS statistical software (version 26.0).

## 3. Results

### 3.1. Study population

Overall, 635 patients were evaluated from October 13, 2020 to December 31, 2021 in this study. After excluding 136 patients, including those with congenital heart disease and ventricular aneurysms and postoperative patients with a pacemaker, the final study population comprised 499 patients (Fig. 2). Due to the lack of tricuspid stenosis cases, we only considered the binary classification problem of the presence or absence of tricuspid stenosis in the tricuspid region. Setting October 1, 2021 as the time cutoff point, patients enrolled before this date were set as the AI algorithm training group and those enrolled after as the validation group. As there were no cases of mitral valve stenosis or aortic valve stenosis after October 1, 2021, three patients with the latest enrollment time were selected from the groups of patients with mitral valve stenosis and aortic valve stenosis before October 1 (a total of six samples) and were included in the result validation group. Therefore, in the AI algorithm training phase, between October 13, 2020 and September 30, 2021, 354 samples were included. In this group, 170 patients with VHD required an intervention. In the results verification phase, 145 samples were included from October 1, 2021 to December 31, 2021. A total of 44 patients in the result validation group had VHD that required further intervention (Table 1 and Fig. 2).

### 3.2. Accuracy of AI is comparable to echocardiography for the diagnosis of VHD

Regarding the diagnostic results, the sensitivity, specificity, and accuracy of VHD ranged from 71.4 to 100.0 %, 83.5–100.0 %, and 84.1–100.0 %, respectively. Among them, the best diagnostic performance was observed in mitral stenosis, with a sensitivity of 100.0 % (31.0–100.0 %), a specificity of 100 % (96.7–100.0 %), and an accuracy of 100 % (97.5–100.0 %). Regarding aortic regurgitation, a sensitivity of 71.4 % (30.3–94.9 %), specificity of 86.2 % (79.1–91.3 %), and accuracy of 85.5 % (78.7–90.8 %) was obtained (Table 2). The results of Accuracy & Loss vs. Epochs of each model are presented in Supplementary Figure S2. The confusion matrices for each VHD in the validation group are presented in Supplementary Table S2.

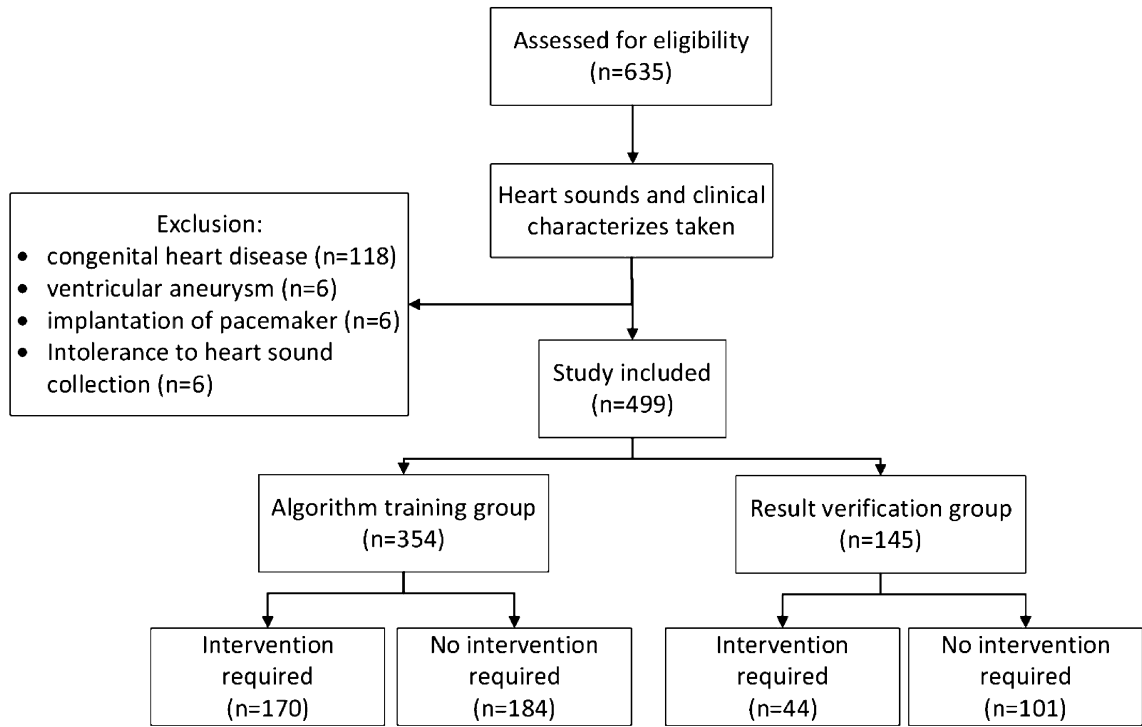


Fig. 2. Flowchart of enrollment.

Table 1  
Baseline characteristics.

	Algorithm training group (n = 354)	Result validation group (n = 145)	P-value
Age (years)	66.00 (57.00, 75.00)	68.00 (61.00, 76.00)	0.001
Sex (male; cases [%])	218 (61.6 %)	70 (48.3 %)	0.006
BMI (kg/m <sup>2</sup> )	24.22 (22.09, 26.83)	24.80 (22.85, 27.51)	0.104
BSA (m <sup>2</sup> )	1.88 (1.72, 2.00)	1.87 (1.73, 1.98)	0.544
Systolic blood pressure (mmHg)	132.40 ± 20.92	132.84 ± 19.60	0.831
Diastolic blood pressure (mmHg)	74.97 ± 12.48	79.17 ± 12.49	0.001
Heart rate (beat per minute)	75.66 ± 15.27	76.29 ± 19.24	0.704
History of coronary heart disease (cases [%])	131 (37.0 %)	113 (77.9 %)	<0.001
History of hypertension (cases [%])	158 (44.6 %)	105 (72.4 %)	<0.001
History of persistent atrial fibrillation (cases [%])	50 (14.1 %)	8 (5.5 %)	0.006
Left atrium anteroposterior diameter (mm)	42.10 (38.00, 48.00)	40.30 (37.95, 46.10)	0.101
Left ventricular diastolic diameter (mm)	51.00 (46.00, 57.00)	48.80 (44.80, 54.15)	0.007
Interventricular septum thickness (mm)	10.00 (9.00, 11.50)	9.10 (8.30, 10.20)	<0.001
Left ventricular ejection fraction (%)	60.00 (56.00, 65.00)	60.00 (50.50, 63.00)	0.017
QRS interval (ms)	98.00 (90.00, 114.00)	99.00 (92.00, 111.50)	0.534
Mixed valve diseases (cases [%])	39 (11.0 %)	5 (3.4 %)	0.007

3.3. Patients with mixed valve diseases were more likely to be misdiagnosed

Next, we performed a multiple logistic regression to identify factors affecting diagnostic consistency. The analysis showed that mixed

Table 2  
Valvular heart disease screening performance.

Valvular heart disease	Sensitivity (95 % CI)	Specificity (95 % CI)	Accuracy (95 % CI)
Mitral stenosis	100.0 % (31.0–100.0 %)	100 % (96.7–100.0 %)	100 % (97.5–100.0 %)
Mitral regurgitation	84.4 % (66.5–94.1 %)	90.3 % (82.9–94.8 %)	89.0 % (82.7–93.6 %)
Tricuspid regurgitation	100.0 % (51.7–100.0 %)	83.5 % (76.0–89.0 %)	84.1 % (77.2–89.7 %)
Aortic stenosis	75.0 % (21.9–98.7 %)	99.3 % (95.5–100.0 %)	98.6 % (95.1–99.8 %)
Aortic regurgitation	71.4 % (30.3–94.9 %)	86.2 % (79.1–91.3 %)	85.5 % (78.7–90.8 %)

valvular lesions were the factor most strongly associated with misdiagnosis (OR 8.17,  $p = 0.064$ ) and were mainly associated with false negatives (OR 72.91,  $p < 0.001$ ). In addition, persistent atrial fibrillation was also associated with false negatives (OR 12.266,  $p = 0.039$ ). No influential factors associated with false positives were found (Supplementary Table S3).

4. Discussion

Our results show that the analysis of raw heart sound data using a deep learning model effectively identifies VHDs requiring further intervention in adults. Using echocardiography as the gold standard for comparison, the sensitivity, specificity, and accuracy ranges of this model were 71.4–100.0 %, 83.5–100.0 %, and 84.1–100.0 %, respectively.

4.1. Raw heart sound data

The salient point of this study is that the model was trained using raw heart sound data without manual processing and annotation as input data, which is more relevant to clinical applications. Previous attempts to automatically classify heart sounds for diagnostic purposes have been



primarily based on unrealistically clear, processed, or manually labeled heart sounds. Some studies evaluated phonocardiograms in their research rather than heart sounds collected directly from a patient's chest wall.[11,14,20,21] The phonocardiogram was filtered for wave and noise reduction processing and did not contain interfering information such as ambient and breath sounds. Furthermore, the acquisition of a phonocardiogram relies on a dedicated cardiograph that records the electrocardiograph synchronously to distinguish the systolic and diastolic phases of the heart; however, cardiographs are not extensively used in clinical practice. Other studies, although using heart sounds as the object of analysis, first required manual annotation to segment the heart sounds and murmurs or to artificially screen the auscultatory region with the largest murmurs for analysis[22,23] and did not fully automate the processing and analysis of heart sounds. In this study, a preprogrammed system automatically performed noise reduction and feature extraction. This automated heart sound processing is similar to the common clinical scenario.

#### 4.2. Outputting a specific diagnosis of VHD

By training the algorithm with echocardiography as the gold standard, the deep learning model used in this study could directly output a specific diagnosis of VHD. This overcomes the limitations of previous studies, which could only output a binary diagnosis of "normal heart sounds" or "abnormal heart sounds." [14,15] For example, the Johns Hopkins Cardiac Auscultatory Recording Database, from which Thompson et al. selected 3,180 recordings (collected from 603 outpatients) for AI analysis, showed that their AI differentiated between normal and abnormal heart sounds with a sensitivity of 93 % (CI 90–95 %), specificity of 81 % (CI 75–85 %), and accuracy of 88 % (CI 85–91 %). These dichotomous results imply that these algorithms can only be used as auscultation aids, not automatic diagnostic tools. Another study used noninvasive wearable inertial sensors to collect heart sounds from 21 patients with aortic stenosis and 13 with a non-aortic stenosis-valvular heart condition. The accuracies of the machine learning algorithms used to identify the heart sounds of patients with aortic stenosis were 0.87 by decision tree, 0.96 by random forest, 0.91 by simple neural network, and 0.95 by XGBoost.[24] However, this study only included a single-site analysis of the aortic valve, whereas our research covered all three valve locations: the aortic, mitral, and tricuspid valves. The output of our study was based on specific VHD diagnoses, such as mitral stenosis and aortic regurgitation. Therefore, these algorithms are likely to provide more focused recommendations for follow-up and can be used to screen for valve disease and guide clinical decisions on a large scale.

#### 4.3. Beyond the echocardiography

This study showed that heart sound analysis using a deep learning model is comparable to echocardiography for VHD diagnosis. Additionally, heart sound analysis surpasses echocardiography in terms of speed and ease of acquisition. First, the acquisition technique can be completed by placing an electronic stethoscope on the patient's chest wall. In addition, body surface positioning of the auscultation zone is simple and does not require medical training. Furthermore, the heart sound acquisition time was short, with each auscultation zone requiring only 30 s. Moreover, the hardware cost of an electronic stethoscope is significantly lower than that of an echocardiographic instrument. These advantages indicate that compared to echocardiography, heart sound screening is a better choice for disease screening in large populations. Additionally, with advancements in digital auscultation hardware, such as the Super StethoScope [25], it is effective as a complementary diagnostic test in primary health facilities, for patient self-screening and follow-up, and in telemedicine. Notably, this technology may reduce the number of patients who must report to the hospital for echocardiography through primary screening, thereby preventing the risk of cross-infection during pandemics, e.g., the coronavirus disease-2019

pandemic. Besides, AI heart sound diagnostic models like ours collect a large amount of heart sound data during training and use. This heart sound data can also be used for physician auscultation training to improve the auscultation skills of physicians and medical students. It is easier to store, playback and teach than echocardiography.

#### 4.4. Limitations

This study had some limitations. First, this study primarily included inpatients admitted to cardiac specialty wards, in whom VHD intervention was more prevalent than in general medical wards or the general population. Therefore, this deep learning model must be validated for screening valve diseases in the general population. Second, a trained researcher auscultated the heart sounds during the cardiac examination; untrained personnel may mislocate the auscultation zone, affecting the patient's self-screening results. The system is designed for primary care or mass population screening; therefore, the operator is expected to have some medical background or training. Additionally, our heart sound collection platform application provides a schematic diagram of the auscultation area to help users locate the auscultation area.[16]

The sensitivity of recognizing aortic valve lesions requiring intervention was 70–75 %, slightly lower than that observed with mitral or tricuspid valve abnormalities; this may be due to a suboptimal correlation between the intensity of the systolic murmur and severity of VHD. Notably, aortic valve disease affects heart function and blood pressure more than other valve diseases, and aortic valve dysfunction impairs the systolic or diastolic function of the left ventricle, decreasing the ejection fraction and murmur intensity.[26] However, if a patient has severe VHD, which usually causes significant clinical symptoms, a false-negative heart sound screen alone may not delay the patient's visit.

#### 4.5. Future research

The deep learning model in this study can be further explored in the future. First, an automatic diagnostic model should be evaluated in different populations, such as patients attending the cardiology outpatient clinic, and used as a large-scale population screening tool. Therefore, we will explore whether this diagnostic model has sufficient diagnostic efficacy in comparison with echocardiography in populations with different rates of VHD prevalence. Second, future attempts may be made to reduce the number of auscultation areas in the patient or to explore whether a heart sound acquisition area different from the conventional sites can achieve a diagnostic efficacy similar to echocardiography. This would reduce the patient sampling time and improve diagnostic efficiency.

In conclusion, based on raw heart sound data, this deep learning model could effectively identify patients with various types of VHD requiring intervention. This model can provide an automatic diagnosis and disease severity assessment of VHD and assist with screening, diagnosis, and follow-up.

##### Funding

This study was funded by the National Natural Science Foundation of China (grant no. U1913210). The funding source had no role in the study design; in the collection, analysis, interpretation of data; in writing the report; or in the decision to submit the article for publication.

##### Ethics approval

All procedures were performed in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and the Helsinki Declaration of 1975, as revised in 2000. Informed consent was obtained from all patients for inclusion in the study. The protocol was approved as a minimal-risk study by the institutional review boards of the participating sites (approval numbers: KY2021K100 and 2020-ZX14).

## CRediT authorship contribution statement

**Zihan Jiang:** Conceptualization, Data curation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing. **Wenhua Song:** Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Yonghong Yan:** Conceptualization, Investigation, Methodology, Supervision. **Ao Li:** Data curation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing. **Yujing Shen:** Data curation, Supervision. **Shouda Lu:** Data curation, Software. **Tonglian Lv:** Data curation. **Xinmu Li:** Data curation. **Ta Li:** Data curation, Formal analysis, Software. **Xueshuai Zhang:** Data curation, Formal analysis, Software. **Xun Wang:** Data curation, Formal analysis, Methodology, Software. **Yingjie Qi:** Data curation. **Wei Hua:** Investigation, Supervision. **Min Tang:** Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. **Tong Liu:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Writing – original draft, 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.

## Acknowledgments

The authors thank Hongda Zhang, MD, and Sixian Weng, MD, from the Arrhythmia Center, State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, and National Center for Cardiovascular Diseases for polishing the article.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijcha.2024.101368>.

## References

- [1] V.T. Nkomo, J.M. Gardin, T.N. Skelton, J.S. Gottdiener, C.G. Scott, M. Enriquez-Sarano, Burden of valvular heart diseases: a population-based study, *Lancet* (London, England). 368 (2006) 1005–1011.
- [2] S. Coffey, B. Cox, M.J. Williams, Lack of progress in valvular heart disease in the pre-transcatheter aortic valve replacement era: increasing deaths and minimal change in mortality rate over the past three decades, *Am Heart J*. 167 (2014) 562–567.e2.
- [3] N. Frey, R.P. Steeds, T.K. Rudolph, et al., Symptoms, disease severity and treatment of adults with a new diagnosis of severe aortic stenosis, *Heart*. 105 (2019) 1709–1716.
- [4] P.A. McCullough, H.S. Mehta, D.P. Cork, et al., The healthcare burden of disease progression in medicare patients with functional mitral regurgitation, *J Med Econ*. 22 (2019) 909–916.
- [5] M. Thoenes, P. Bramlage, P. Zamorano, et al., Patient screening for early detection of aortic stenosis (AS)-review of current practice and future perspectives, *J Thorac Dis* 10 (2018) 5584–5594.
- [6] S.K.M. Gardezi, S.G. Myerson, J. Chambers, et al., Cardiac auscultation poorly predicts the presence of valvular heart disease in asymptomatic primary care patients, *Heart*. 104 (2018) 1832–1835.
- [7] J.M. Vukanovic-Criley, S. Criley, C.M. Warde, et al., Competency in cardiac examination skills in medical students, trainees, physicians, and faculty: a multicenter study, *Arch Intern Med*. 166 (2006) 610–616.
- [8] A. Raza, A. Mehmood, S. Ullah, M. Ahmad, G.S. Choi, B.W. On, Heartbeat sound signal classification using deep learning, *Sensors* (basel) 19 (2019) 4819.
- [9] W. Han, Z. Yang, J. Lu, S. Xie, Supervised threshold-based heart sound classification algorithm, *Physiol Meas*. 39 (2018) 115011.
- [10] B. Bozkurt, I. Germanakis, Y. Stylianou, A study of time-frequency features for CNN-based automatic heart sound classification for pathology detection, *Comput Biol Med*. 100 (2018) 132–143.
- [11] S.L. Oh, V. Jahmunah, C.P. Ooi, et al., Classification of heart sound signals using a novel deep WaveNet model, *Comput Methods Programs Biomed*. 196 (2020) 105604.
- [12] C. Mitchell, P.S. Rahko, L.A. Blauwet, et al., Guidelines for performing a comprehensive transthoracic echocardiographic examination in adults: recommendations from the American Society of Echocardiography, *J Am Soc Echocardiogr*. 32 (2019) 1–64.
- [13] C.M. Otto, R.A. Nishimura, R.O. Bonow, et al., 2020 ACC/AHA guideline for the Management of Patients with Valvular Heart Disease: a report of the American College of Cardiology/American Heart Association joint committee on clinical practice guidelines, *J Am Coll Cardiol*. 77 (2021) e25–e197.
- [14] W.R. Thompson, A.J. Reinisch, M.J. Unterberger, A.J. Schrieffer, Artificial intelligence-assisted auscultation of heart murmurs: validation by virtual clinical trial, *Pediatr Cardiol*. 40 (2019) 623–629.
- [15] G.D. Clifford, C. Liu, B. Moody, et al., Recent advances in heart sound analysis, *Physiol Meas*. 38 (2017) E10–E25.
- [16] S. Yujing, W. Xun, T. Min, Design and implementation of heart sound acquisition application software on android mobile platform, *Chin Med Equip J*. 041 (2020) 38–43, <https://doi.org/10.19745/j.1003-8868.2020034>.
- [17] S. Yujing, W. Xun, T. Min, Combination of optimally-modified log-spectral amplitude estimator and wavelet for heart sound denoising, *Chin J Med Physics*. 37 (2020) 1287–1292, <https://doi.org/10.3969/j.issn.1005-202X.2020.10.013>.
- [18] P.K. Jain, A.K. Tiwari, An adaptive thresholding method for the wavelet based denoising of phonocardiogram signal, *Biomed Signal Process Control*. 38 (2017) 388–399.
- [19] R.G. Newcombe, Two-sided confidence intervals for the single proportion: comparison of seven methods, *Stat Med*. 17 (1998) 857–872.
- [20] M. Sotaquirá, D. Alvear, M. Mondragón, Phonocardiogram classification using deep neural networks and weighted probability comparisons, *J Med Eng Technol*. 42 (2018) 510–517.
- [21] C. Potes S. Parvaneh A. Rahman B. Conroy Ensemble of Feature-based and Deep learning-based Classifiers for Detection of Abnormal Heart Sounds. In, computing in cardiology conference (CinC) IEEE. 2016 2016 621 624.
- [22] J.S. Chorbha, A.M. Shapiro, L. Le, et al., Deep learning algorithm for automated cardiac murmur detection via a digital stethoscope platform, *J Am Heart Assoc*. 10 (2021) e019905.
- [23] L.S. Lai, A.N. Redington, A.J. Reinisch, M.J. Unterberger, A.J. Schrieffer, Computerized automatic diagnosis of innocent and pathologic murmurs in pediatrics: a pilot study, *Congenit Heart Dis*. 11 (2016) 386–395.
- [24] C. Yang, B.D. Ojha, N.D. Aranoff, P. Green, N. Tavassolian, Classification of aortic stenosis using conventional machine learning and deep learning methods based on multi-dimensional cardio-mechanical signals, *Sci Rep*. 10 (2020) 17521.
- [25] S. Ogawa, F. Namino, T. Mori, G. Sato, T. Yamakawa, S. Saito, AI diagnosis of heart sounds differentiated with super StethoScope, *J Cardiol*. S0914–S087 (23) (2023) 00229.
- [26] P. Das, C. Pocock, J. Chambers, The patient with a systolic murmur: severe aortic stenosis may be missed during cardiovascular examination, *QJM* 93 (2000) 685–688.