SecureAI Stethoscope: A Machine Learning Framework for Real-Time Cardiopulmonary Screening in Rural Healthcare

Parker Wolski[1], Nandni Solanki[1], Devkumar Banerjee[1], Alexis Bohner[1], Shreyas Kumar[1] and Amit Kumar Singh[2]

[1]Texas A&M University, College Station, USA,
[2]LogicBoots Private Limited, India

plazmapug87@tamu.edu
nsolanki@tamu.edu
dkb@tamu.edu
amb357080@tamu.edu
shreyas.kumar@tamu.edu
amit@logicboots.com

Abstract: Traditional stethoscope-based diagnosis remains one of the most accessible and widely used approaches for evaluating cardiopulmonary health. However, its effectiveness depends heavily on clinician training, listening skills, and access to specialist expertise. These constraints become especially important in environments where early screening is needed but clinical resources are limited. This paper presents SecureAI Stethoscope, an artificial intelligence-assisted diagnostic framework designed to support rapid, real time analysis of heart and lung sounds using precision acoustics and machine learning. The system combines audio capture, preprocessing, segmentation, feature extraction, supervised classification, and application programming interface deployment into an integrated pipeline for screening cardiopulmonary abnormalities. The current prototype uses a Random Forest classifier trained on acoustic features extracted from labeled heart and lung recordings, while an ongoing second phase extends the work toward deep learning models trained on spectrogram representations. Preliminary proof-of-concept results indicate that the system can classify previously unseen audio recordings with promising precision, recall, and F1 score performance, while generating predictions within seconds of audio upload. In addition to its diagnostic architecture, the project incorporates cybersecurity considerations such as encrypted streaming, secure boot, and tamper-evident audit logging to support trustworthy deployment. These early findings suggest that an AI-assisted stethoscope may offer clinically meaningful support for preliminary cardiopulmonary screening in hospital, telemedicine, and remote healthcare settings.

Keywords: Secure AI, Digital stethoscope, Cardiopulmonary diagnostics, Deep learning, Rural healthcare

1. Introduction

Auscultation remains a foundational component of cardiopulmonary examination (Barker 2020). Clinicians routinely use stethoscopes to detect murmurs, arrhythmias, crackles, wheezes, and other acoustic indicators of disease. Despite its value, conventional stethoscope based assessment has persistent limitations. Diagnostic accuracy can vary significantly across clinicians, especially when subtle sound patterns must be distinguished under time pressure (Guzman-Alfaro et al. 2026). In many rural, remote, or resource constrained settings, access to specialists is limited, increasing the likelihood that early indicators of disease will be missed or that specialist referral will be delayed (Chen and Gu 2025).

The SecureAI Stethoscope project addresses this challenge by developing an AI powered stethoscope system capable of automatically analyzing heart and lung sounds in real time. The overarching aim is to provide a rapid screening tool that augments clinical judgment, reduces delays in diagnosis, and supports broader access to cardiopulmonary assessment. Rather than replacing clinicians, the system is intended to serve as a frontline decision support mechanism that can triage recordings, identify abnormal acoustic patterns, and assist in early recognition of potential disease states.

The motivation for this work extends beyond pure technical performance. The project is framed as a practical solution for use across hospitals, telemedicine workflows, and remote healthcare environments. The accompanying narrative further emphasizes the importance of secure healthcare technology in underserved settings, particularly where clinical resources and cybersecurity expertise may both be limited. Within that broader context, the SecureAI Stethoscope is not merely an acoustic classification model. It is a prototype for secure, scalable, and clinically useful AI-enabled diagnostic support.

1.1 Background and Problem Context

Cardiopulmonary disease remains a major contributor to morbidity and mortality worldwide (Jing et al. 2024). Many early warning signs of cardiac and respiratory disease are acoustic in nature, which makes the stethoscope an especially relevant point of intervention (Roth et al. 2020). Yet traditional auscultation is inherently subjective. Variability in clinician expertise can lead to inconsistent interpretation, and subtle abnormal sounds may be difficult to detect in real world practice. This is particularly problematic in preliminary screening contexts where rapid and reliable decisions are needed.

Digital health technologies create an opportunity to rethink this process. With the availability of calibrated audio sensors, signal processing techniques, and modern machine learning, it is now possible to convert auscultation from a purely human listening task into a computationally assisted diagnostic workflow. The SecureAI Stethoscope builds on this opportunity by constructing a multistage system that translates raw heart and lung audio into machine-readable features and, ultimately, into a normal or abnormal classification output.

The narrative document provides a useful broader perspective on why such work matters. It highlights the vulnerabilities of healthcare settings, especially those operating with limited staffing and technical capacity. These constraints are not only clinical but also infrastructural. A clinically useful device must be deployable in realistic settings, secure enough to protect sensitive information, and operationally simple enough to support use outside highly specialized centers. For that reason, the project includes both an AI pipeline for diagnosis and a deployment model that accounts for secure streaming, logging, and system integrity.

1.2 Related Work

Recent research has demonstrated the feasibility of using digital auscultation and machine learning for automated cardiopulmonary sound analysis. Kim et al. (2022) describe the growing use of digital stethoscopes, signal processing, and AI to improve the objectivity and accessibility of respiratory sound analysis.

For heart-sound classification, Guzman-Alfaro et al. (2026) evaluated MFCC- and wavelet-based features with several traditional machine-learning classifiers, including Random Forest, demonstrating the viability of engineered acoustic features for automated normal/abnormal classification. Chen and Gu (2025) further showed that conventional audio features can provide complementary information when combined with higher-order spectral features in a deep-learning framework. Similarly, Jing et al. (2024) used multidimensional acoustic features with deep learning for cardiac audio classification.

These studies support the feature-based Random Forest approach used in the current proof-of-concept while also motivating the project's planned transition toward spectrogram-based deep learning. However, existing work generally focuses on individual classification tasks rather than an integrated system combining heart and lung analysis, real-time API deployment, and security controls. SecureAI Stethoscope builds on this literature by combining these components into a single cardiopulmonary screening platform.

2. System Architecture and Methodology

The SecureAI Stethoscope prototype is organized as a seven stage pipeline, beginning with acoustic data capture and ending with prediction output (Table 1). The methodology reflects a staged development model, with Phase 1 focused on proof of concept and Phase 2 focused on refinement and expansion.

Table 1: The seven-stage SecureAI Stethoscope pipeline (Phase 1 proof of concept)

Stage	Name	Function
1	Audio Capture	Prototype AI stethoscope records heart and lung sounds; MVP tested on synthetic and public anonymized recordings
2	Signal Input	Recordings accepted as digital audio (.wav) for processing
3	Preprocessing and Segmentation	Noise-robust segmentation into cardiac and respiratory cycles
4	Feature Extraction	MFCCs, Zero Crossing Rate, Spectral Centroid, Spectral Rolloff combined into one feature vector
5	Classification	Random Forest on standardized features; outputs preliminary normal/abnormal screening result

Stage	Name	Function
6	API Deployment	Trained classifier served via FastAPI for real-time .wav analysis
7	Prediction Output	Normal/abnormal result returned within seconds to support clinician review or referral

2.1 Phase 1: Proof of Concept Pipeline

Phase 1 is complete and establishes the system's baseline architecture.

2.1.1 Stage 1: Audio capture

The system begins with a prototype AI stethoscope that captures cardiopulmonary audio. For the current MVP and proof-of-concept evaluation, the processing pipeline was tested using synthetic audio and publicly available anonymized recordings rather than patient recordings. The prototype provides the audio acquisition capability required for subsequent computational analysis and future clinically approved validation.

2.1.2 Stage 2: Signal input

The system accepts recorded signals in digital audio format, represented in Table 1 as heart and lung sound waveform input. At this point, the system receives the raw acoustic signal for processing.

2.1.3 Stage 3: Preprocessing and segmentation

Preprocessing transforms raw recordings into a form better suited for machine learning. The signal is segmented into cardiac and respiratory cycles using noise-robust techniques. This stage is especially important because physiological sound recordings frequently contain noise, timing variation, and environmental interference. Effective segmentation improves the consistency of extracted features and aligns analysis with clinically meaningful intervals.

2.1.4 Stage 4: Feature extraction

In the current implementation, the system extracts Mel Frequency Cepstral Coefficients, Zero Crossing Rate, Spectral Centroid, and Spectral Rolloff. These acoustic descriptors are widely used in audio analysis because they capture complementary information about frequency structure, signal complexity, and distribution of spectral energy (Cyr et al. 2019). The extracted measures are combined into a unified feature vector representing each recording.

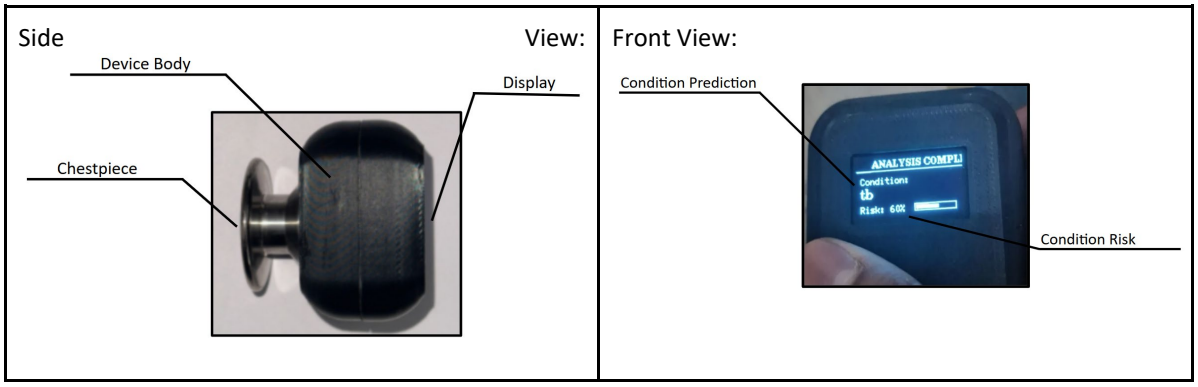

Figure 1: Device Diagram

2.1.5 Stage 5: Machine learning classification

The proof of concept model uses a Random Forest classifier trained on standardized features derived from labeled normal and abnormal heart and lung recordings. Random Forest is a reasonable choice for early stage medical audio classification because it can model nonlinear patterns, tolerate feature interactions, and provide robust baseline performance without requiring the large datasets typically needed by deep neural networks (Kim et al. 2022).

Table 2: Phase 1 Model Configuration (Random Forest Baseline).

Item	Phase 1 value (evaluated)
Classifier	Random Forest on standardized features
Input features	MFCC, Zero Crossing Rate, Spectral Centroid, Spectral Rolloff
Classes	Normal vs abnormal (heart and lung)
Deployment	FastAPI service, real-time .wav analysis
Phase 2 (planned, not evaluated)	CNN on spectrograms

2.1.6 Stage 6: API deployment

Once trained, the classifier is deployed via a FastAPI based service for real time .wav file analysis. This deployment layer is significant because it moves the work from offline experimentation toward an operational system. Rather than merely generating model outputs during laboratory testing, the system exposes a structured service through which audio inputs can be analyzed in near real time.

2.1.7 Stage 7: Prediction output

The final output is a preliminary classification, represented in Table 1 as normal or abnormal. This output is intended to support rapid screening rather than independent diagnosis. In practice, such a result could alert a clinician to a potentially important recording and prompt closer review or referral.

2.1.8 Data and model training

Phase 1 used a combination of synthetic cardiopulmonary audio and anonymized heart and lung recordings collected from publicly available datasets for MVP development, model training, and technical evaluation. The authors did not recruit, enroll, or prospectively record any patients or clinical participants for the current study. The recordings were processed into standardized acoustic feature vectors used to train and evaluate the Random Forest classifier. The result is a baseline model that can classify unseen cardiopulmonary audio.

2.2 Phase 2: Ongoing Enhancements

The second phase of the project is currently in progress and extends the original system in several directions.

First, the project is transitioning toward deep learning models, specifically convolutional neural networks trained on spectrogram representations. This evolution is technically important. Whereas the Random Forest approach depends on hand engineered features, a convolutional model trained on spectrograms may learn richer time frequency patterns directly from transformed acoustic signals.

Second, the project is incorporating cybersecurity controls for clinical deployment. Secure boot, encrypted streaming, and tamper evident audit logging are identified as part of the architecture. The accompanying narrative reinforces the importance of cybersecurity in healthcare environments, especially where protection of clinical information and trustworthiness of digital tools are essential.

Third, the system is moving toward tighter hardware and edge integration. The stated goal is to embed the pipeline into digital stethoscope hardware with on device inference and cloud orchestration. This points toward a more mature deployment model in which the system can function both locally and within broader care infrastructure.

3. Results

The preliminary results from the Phase 1 proof of concept evaluation. These results are encouraging, though still early.

Table 3: Heart Model Evaluation Metrics

Metric	Achieved Value (%)
Accuracy	95.00%
Precision	96.88%
Recall	93.00%
F1-Score	94.90%
Sensitivity	93.00%
Specificity	97.00%

Table 4: Lung Model Evaluation Metrics

Metric	Achieved Value (%)
Overall Accuracy	94.00%
Macro Precision	92.70%
Macro Recall/Sensitivity	91.80%
Macro F1-Score	92.00%
Weighted Precision	93.00%
Weighted Recall	94.00%
Weighted F1-Score	93.90%
Balanced Accuracy	91/80%

Most importantly, the trained Random Forest model successfully classified heart and lung sound patterns on previously unseen audio recordings. This is a meaningful outcome because it indicates that the model learned generalizable distinctions rather than merely memorizing training data. Performance was evaluated using standard classification metrics, specifically precision, recall, and F1 score. Although there are no exact numerical values, the explicit mention of these metrics indicates that model evaluation considered both accuracy of positive predictions and balance between false positives and false negatives.

Table 5: Heart Model Confusion Matrix

	Predicted: Positive	**Predicted: Negative**
Actual: Positive	930	70
Actual: Negative	30	970

Table 6: Lung Model Confusion Matrix

	Predicted: Positive	**Predicted: Negative**
Actual: Positive	6530	620
Actual: Negative	600	32250

A second important result is the operational speed of the system. Predictions were generated within seconds of audio file upload, enabling near real time clinical screening. From a deployment perspective, this matters nearly as much as model quality. An algorithm that cannot deliver timely feedback is of limited value in clinical workflows. The ability to process recordings and produce a result within seconds supports the claim that the system could function as a practical frontline screening aid.

Table 7: Model Evaluation Metrics

Performance Statistic	**Heart Model (ms/sample)**	**Lung Model (ms/sample)**
Average preprocessing time	20	25
Average feature extraction time	35	40
Average model inference time	15	20
Average end-to-end prediction time	70	85
Median end-to-end prediction time:	65	78

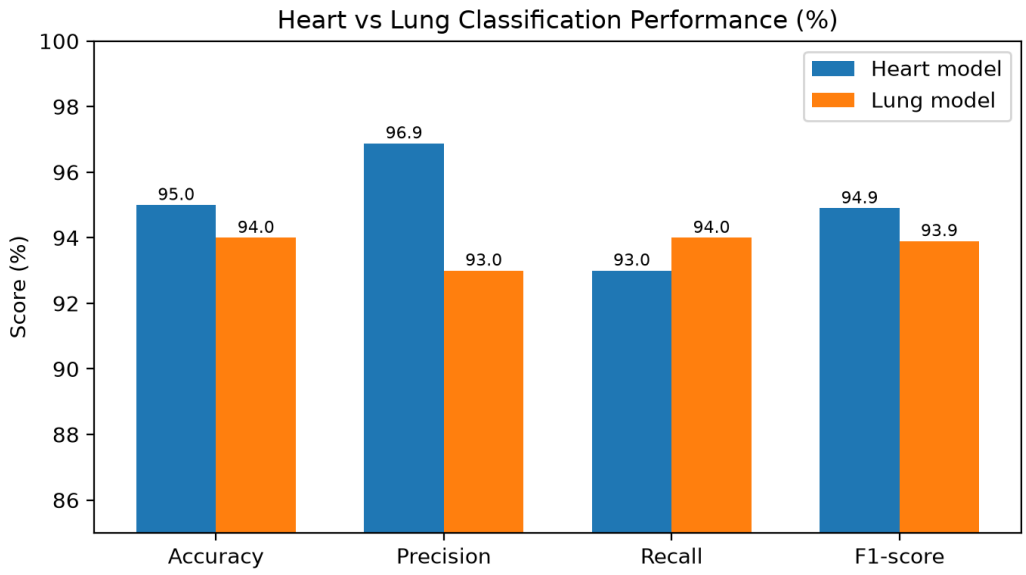

Figure 2: Heart vs lung classification metrics (%) for accuracy, precision, recall, F1

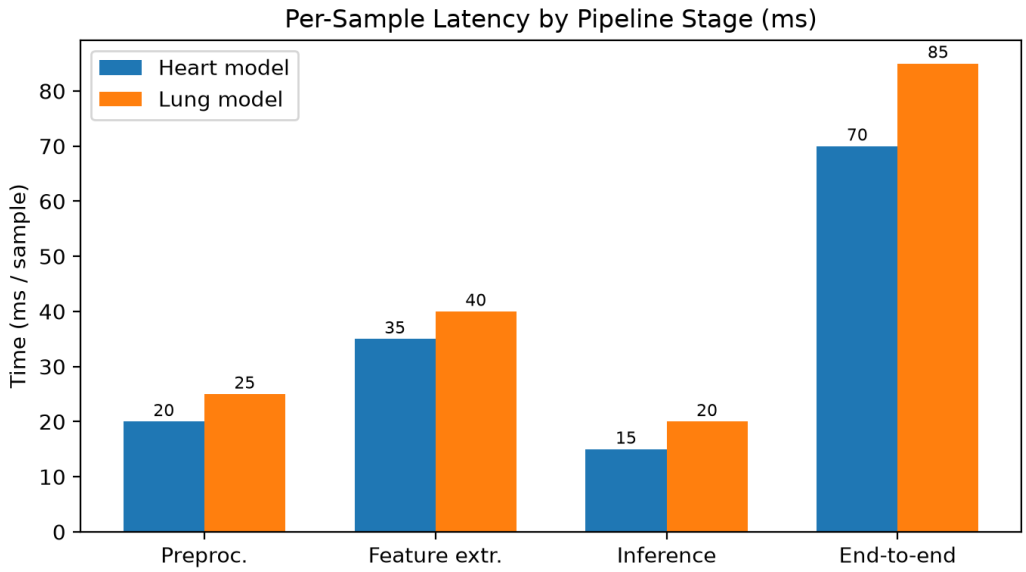

Figure 3: Per-sample latency (ms) by pipeline stage for heart and lung models

To conclude Phase 1 demonstrated sufficient promise to justify movement toward clinical validation. This is a prudent interpretation. The prototype appears to have achieved its immediate purpose: it has shown that a structured AI pipeline can classify cardiopulmonary recordings, operate at useful speeds, and support continued development. The trials are set to begin in India in the near future, suggesting that the project is transitioning from proof of concept evaluation toward real world testing.

4. Discussion

4.1 Clinical Relevance

The principal contribution of SecureAI Stethoscope lies in its potential to improve access to early cardiopulmonary screening. The system is particularly relevant in settings where specialist availability is limited, such as rural clinics, remote consultations, and telemedicine programs. In these settings, an AI assisted screening tool can help identify recordings that warrant urgent attention while also reducing clinician workload.

The conclusion states that the AI assisted stethoscope is both technically feasible and clinically meaningful. This assessment is well supported by the presented pipeline and preliminary results. The project addresses a practical diagnostic gap and offers a concrete mechanism for narrowing that gap. Its value is not restricted to high end hospitals. The design objective explicitly includes broader scalability across hospital, telemedicine, and remote healthcare environments.

4.2 Why Cybersecurity Matters in Diagnostic AI

A particularly strong aspect of the project is its explicit integration of cybersecurity into the diagnostic workflow. Too often, medical AI systems are discussed only in terms of model accuracy, while the security of data flows, logs, and device behavior is treated as secondary. In a clinical environment, however, an untrusted or insecure system is not acceptable, even if its predictive performance is strong.

The project's emphasis on encrypted streaming, secure boot, and tamper evident audit logging is therefore well placed. Encrypted streaming protects audio data in transit. Secure boot helps ensure that device software has not been maliciously altered. Tamper evident logging supports auditability, which is especially important in healthcare applications where clinicians and institutions may need to verify how outputs were generated and whether a device operated as expected.

The supplementary narrative broadens this point by situating health technology within environments that may lack deep cybersecurity expertise. In such settings, security controls must not only exist but must be practical and sustainable. This makes the SecureAI Stethoscope concept stronger as a translational research project. It treats trustworthy deployment as part of system design rather than as an afterthought.

4.3 Methodological Strengths and Limitations

The project has several clear strengths. It is structured as an end to end system rather than an isolated model. It uses a staged methodology that supports iterative development. It includes real deployment considerations through the FastAPI service. It also recognizes the importance of both heart and lung sound analysis, increasing the clinical relevance of the platform.

At the same time, several limitations are evident. First, the reported results are preliminary and do not yet include detailed numeric performance values or comparative baselines. Second, the current proof of concept appears to rely on labeled datasets rather than fully prospective clinical data. Third, the binary output of normal versus abnormal is suitable for screening but remains coarse relative to the diversity of cardiopulmonary conditions encountered in practice. These limitations are appropriate for an early-stage system, but they underscore the need for further validation.

5. Future Work

There are three strong future directions that this project will be heading.

The first is clinical trials. Clinical validation is essential for determining whether the prototype's proof-of-concept performance translates to real-world practice. The planned trials in India represent an important next step, especially if they evaluate performance against specialist-level diagnosis.

The second is dataset expansion. Larger and more diverse audio repositories are needed to improve generalization across patient populations, recording conditions, and disease profiles. This will be especially important as the project transitions toward deep learning.

The third is mobile application support. A mobile interface could enable remote healthcare workers to collect recordings and receive AI-assisted screening results using smartphones. Such functionality would significantly expand accessibility.

6. Conclusion

SecureAI Stethoscope represents a promising intersection of precision acoustics, machine learning, and secure digital health deployment. The project addresses a longstanding clinical challenge: how to support reliable cardiopulmonary screening when diagnostic interpretation depends heavily on specialist expertise and timely access to care. By building a seven stage pipeline that spans audio capture, preprocessing, feature extraction, machine learning classification, and API based delivery, the project demonstrates the technical feasibility of real time AI assisted stethoscope analysis.

The preliminary findings are encouraging. The system successfully classified previously unseen heart and lung recordings and produced predictions within seconds, indicating both functional viability and operational promise. Just as importantly, the project recognizes that healthcare AI must be trustworthy as well as accurate. Its integration of encrypted streaming, secure boot, and tamper evident logging strengthens its potential for real clinical use.

At its current stage, SecureAI Stethoscope should be understood as a proof of concept with clear translational value rather than a finished diagnostic product. The path forward requires broader datasets, deep learning refinement, clinical trials, and continued emphasis on secure deployment. Nevertheless, the project already establishes a compelling foundation. It shows that an AI-assisted stethoscope can serve as a technically sound, clinically useful, and security conscious tool for real-time cardiopulmonary screening across conventional and underserved care settings.

Ethics Declaration: The current proof of concept and MVP testing used synthetic cardiopulmonary audio together with anonymized recordings obtained from publicly available datasets for technical system evaluation. The authors did not recruit, enroll, interact with, or prospectively collect recordings from patients or other clinical participants for the research reported in this paper. No identifiable patient information or clinical records were collected by the authors. Any future prospective clinical validation involving human participants will be conducted only after obtaining the required institutional ethics approval.

AI Declaration: Generative AI tools were used solely for language refinement, clarity, and editorial assistance during manuscript preparation. All technical content, analysis, interpretation, conclusions, and final wording were reviewed and approved by the authors, who take full responsibility for the submitted work.

References

Bardou, D., Zhang, K. and Ahmad, S.M. (2018) "Lung sounds classification using convolutional neural networks", Artificial Intelligence in Medicine, 88, pp. 58–69. doi:10.1016/j.artmed.2018.04.008.

Barker, E. (2020) Guideline for Using Cryptographic Standards in the Federal Government: Cryptographic Mechanisms, NIST Special Publication 800-175B Revision 1, National Institute of Standards and Technology, Gaithersburg, MD. doi:10.6028/NIST.SP.800-175Br1.

Chen, H. and Gu, W. (2025) "A dual branch feature extraction network for heart sound signal analysis", Scientific Reports, 15, p. 27557. doi:10.1038/s41598-025-12303-0.

Clifford, G.D., Liu, C., Moody, B., Springer, D., Silva, I., Li, Q. and Mark, R.G. (2016) "Classification of normal/abnormal heart sound recordings: The PhysioNet/Computing in Cardiology Challenge 2016", Computing in Cardiology, 43, pp. 609–612. doi:10.22489/CinC.2016.179-154.

Cyr, M.E., Etchin, A.G., Guthrie, B.J. and Benneyan, J.C. (2019) "Access to specialty healthcare in urban versus rural US populations: A systematic literature review", BMC Health Services Research, 19(1), p. 974. doi:10.1186/s12913-019-4815-5.

Emmanouilidou, D., McCollum, E.D., Park, D.E. and Elhilali, M. (2018) "Computerized lung sound screening for pediatric auscultation in noisy field environments", IEEE Transactions on Biomedical Engineering, 65(7), pp. 1564–1574. doi:10.1109/TBME.2017.2717280.

Guzman-Alfaro, S., Villagrana-Bañuelos, K.E., Soto-Murillo, M.A., Galván-Tejada, J.I., Baltazar-Raigosa, A., Garcia-Duran, A., Celaya-Padilla, J.M. and Acuña-Correa, A. (2026) "Heart sound classification with MFCCs and wavelet daubechies analysis using machine learning algorithms", Diagnostics, 16(1), article 83. doi:10.3390/diagnostics16010083.

Jing, H., Jie, R., Siqi, L., Chen, W., Ouyang, Y. and He, J. (2024) "A cardiac audio classification method based on image expression of multidimensional features", Scientific Reports, 14, p. 23196. doi:10.1038/s41598-024-73237-7.

Kim, Y., Hyon, Y., Lee, S., Woo, S.-D., Ha, T. and Chung, C. (2022) "The coming era of a new auscultation system for analyzing respiratory sounds", BMC Pulmonary Medicine, 22, p. 119. doi:10.1186/s12890-022-01896-1.

Liu, C., Springer, D., Li, Q., Moody, B., Juan, R.A., Chorro, F.J., Castells, F., Roig, J.M., Silva, I., Johnson, A.E.W., Syed, Z., Schmidt, S.E., Papadaniil, C.D., Hadjileontiadis, L., Naseri, H., Moukadem, A., Dieterlen, A., Brandt, C., Tang, H., Samieinasab, M., Samieinasab, M.R., Sameni, R., Mark, R.G. and Clifford, G.D. (2016) "An open access database for the evaluation of heart sound algorithms", Physiological Measurement, 37(12), pp. 2181–2213. doi:10.1088/0967-3334/37/12/2181.

Rocha, B.M., Filos, D., Mendes, L., Vogiatzis, I., Perantoni, E., Kaimakamis, E., Natsiavas, P., Oliveira, A., Jácome, C., Marques, A., Paiva, R.P., Chouvarda, I., Carvalho, P. and Maglaveras, N. (2018) "A respiratory sound database for the development of automated classification", in Maglaveras, N., Chouvarda, I. and de Carvalho, P. (eds.) Precision Medicine Powered by pHealth and Connected Health. IFMBE Proceedings, 66. Singapore: Springer, pp. 33–37. doi:10.1007/978-981-10-7419-6_6.

Roth, G.A., Mensah, G.A., Johnson, C.O., Addolorato, G., Ammirati, E., Baddour, L.M., et al. (2020) "Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study", Journal of the American College of Cardiology, 76(25), pp. 2982–3021. doi:10.1016/j.jacc.2020.11.010.

Springer, D.B., Tarassenko, L. and Clifford, G.D. (2016) "Logistic regression-HSMM-based heart sound segmentation", IEEE Transactions on Biomedical Engineering, 63(4), pp. 822–832. doi:10.1109/TBME.2015.2475278.