Journal of
Biomedical Engineering and Research

Review Article Open Access

Development of an Artificial Intelligence-Enhanced Digital Stethoscope: Key Elements for Engineers and Physicians

Emmanuel Andrès[1,2*], Amir Hajjam El Hassani[3] and Noël Lorenzo-Villalba[1,2]

[1]Department of Internal Medicine, Hôpitaux Universitaires de Strasbourg (HUS), Strasbourg, France
[2]EA 3072, Université de Strasbourg, Strasbourg, France
[3]UTBM — Université de Technologie de Belfort-Montbéliard, Belfort, France

[*]**Corresponding Author:** Emmanuel Andrès, Department of Internal Medicine, Hôpitaux Universitaires de Strasbourg (HUS), Strasbourg, France

Received Date: June 16, 2026 **Accepted Date:** June 29, 2026 **Published Date:** July 02, 2026

Citation: Emmanuel Andrès, Amir Hajjam El Hassani, Noël Lorenzo-Villalba (2026) Development of an Artificial Intelligence-Enhanced Digital Stethoscope: Key Elements for Engineers and Physicians. J Biomed Eng Res 10: 1-12

Abstract

The stethoscope, one of the most emblematic instruments in the history of medicine, is undergoing a fundamental transformation. Artificial intelligence-enhanced digital stethoscopes are evolving into intelligent biomedical platforms that combine high fidelity acoustic sensing, real-time embedded signal processing, and deep learning algorithms capable of classifying cardiopulmonary sounds with clinically actionable accuracy. Rather than replacing bedside medicine, these devices aim to revitalize the clinical examination by merging traditional medical observation with advanced computational intelligence. This narrative review synthesizes the key technical and clinical elements required for the design, validation, and deployment of such devices, targeting an audience of both biomedical engineers and internists. We address acoustic physics, sensor technologies, signal processing pipelines, artificial intelligence architectures, regulatory frameworks, telemedicine integration through the E-care platform, and clinical use cases spanning cardiology, pulmonology, and primary care. Fifty peer-reviewed references, all-verifiable on *PubMed* or in primary regulatory sources, underpin the evidence presented.

Keywords: Digital Stethoscope; Artificial Intelligence; Deep Learning; Auscultation; Cardiac Sounds; Lung Sounds; Telemedicine; MEMS; Signal Processing; E-care

©2026 The Authors. Published by the JScholar under the terms of the Crea-tive Commons Attribution License http://creativecommons.org/licenses/by/3.0/, which permits unrestricted use, provided the original author and source are credited.

1. Introduction

The stethoscope remains one of the most emblematic instruments in medicine. Since René Laennec introduced mediate auscultation in 1816, it has occupied a central role both in bedside diagnosis and in the symbolic representation of the medical profession. Yet traditional acoustic auscultation carries well-documented limitations: its interpretation depends heavily on clinician experience, acoustic environments are frequently suboptimal, inter-observer variability is substantial, and the ephemeral nature of the acoustic signal precludes storage, transmission, or algorithmic analysis [1,2].

These limitations have long stimulated the development of digital stethoscopes capable of amplifying, recording, filtering, and transmitting cardiopulmonary sounds. The integration of artificial intelligence into these systems now constitutes a major technological transition. Contemporary AI-enhanced stethoscopes are converging into intelligent biomedical platforms that combine acoustic sensing, embedded electronics, signal processing, cloud computing, and machine learning — systems capable of supporting automated detection of murmurs, respiratory abnormalities, arrhythmias, and other clinically relevant acoustic signatures [1-3].

The development of these technologies requires close collaboration among physicians, biomedical engineers, acousticians, computer scientists, and regulatory experts. The ASAP (*Analyse des Sons Auscultatoires Pathologiques*) project — a collaboration between the *Hôpitaux Universitaires de Strasbourg*, the *Université de Technologie de Belfort-Montbéliard* and *Alcatel* (France) — illustrates this multidisciplinary imperative. The stethoscope prototype (*eStetho*) validation study (n = 857 patients) demonstrated ~94% concordance with the *Littmann Classic III* reference stethoscope and a Cohen's kappa of approximately 0.87 across the full spectrum of internal medicine auscultatory findings [4,5].

The present review provides a structured and practically oriented reference for multidisciplinary teams engaged in the design, clinical evaluation, or deployment of AI-enhanced digital stethoscopes. It covers acoustic foundations, sensor and hardware engineering, signal processing, AI model design, clinical validation, regulatory pathways, and telemedicine integration.

2. Acoustic Foundations of Cardiopulmonary Auscultation

2.1 Cardiac Sounds

Heart sounds are biomechanical vibrations generated by valvular closure, ventricular contraction, and turbulent blood flow. The normal first (S1) and second (S2) heart sounds are predominantly low-frequency events, with dominant energy between 20 and 150 Hz and clinically relevant harmonics extending to approximately 600 Hz. Pathological murmurs may extend beyond 600 Hz; their timing (systolic, diastolic, or continuous), morphology (crescendo-decrescendo, plateau, or rumble), intensity grading (Levine scale I–VI), and radiation pattern encode diagnostic information that an AI classifier must disentangle from background noise [6,7].

Third (S3) and fourth (S4) sounds, gallop rhythms, the opening snap of mitral stenosis, and the pericardial friction rub each possess distinct spectrotemporal signatures. The pericardial rub — a to-and-fro leathery sound — illustrates the challenge of differentiating true pathological findings from artefacts generated by chest hair or clothing movement, a concrete problem for robust algorithm design. Thoracic tissues attenuate acoustic transmission non-uniformly, and engineers must optimize chest piece geometry, mechanical coupling, and transducer placement to preserve diagnostically relevant information [8].

2.2 Respiratory Sounds

Respiratory sounds arise from airflow within the bronchial tree and from interactions between airways and pulmonary tissues. Normal vesicular breath sounds are generated by turbulent airflow in large airways and attenuate progressively above 500 Hz as they traverse lung parenchyma and chest wall. Their acoustic spectrum extends from approximately 100 to 1,000 Hz in healthy adults, but thoracic configuration, adiposity, and the presence of pleural or parenchymal disease alter transmission substantially [9,10].

Adventitious sounds are the primary targets for automated detection. Fine inspiratory crackles, characteristic of pulmonary fibrosis and early pulmonary edema, are brief (<10 ms), high frequency (500–2,000 Hz) discontinuous transients; coarse crackles in bronchiectasis are longer (10–50 ms) and lower in pitch. Wheezes are musical, continuous adventitious sounds generated by airway narrowing in asthma and COPD; rhonchi reflect secretion-induced airway turbulence. The standardized nomenclature established by the *International Lung Sounds Association* (ILSA) provides the essential taxonomic reference for training dataset annotation and inter-study comparison [11-13].

3. Sensor Technologies

3.1 Piezoelectric and MEMS Transducers

Piezoelectric contact microphones were among the earliest sensors used in electronic auscultation. By converting mechanical thoracic vibrations directly into electrical signals, they offer excellent rejection of ambient acoustic noise and good sensitivity for low-frequency cardiac sounds. However, their non-flat frequency response — with resonance peaks that vary with coupling pressure and chest piece mass — requires careful calibration and limits high--frequency fidelity [14,15].

Microelectromechanical systems (MEMS) condenser microphones have progressively become dominant in contemporary digital stethoscopes. They achieve signal--to-noise ratios exceeding 65 dB at 1 Pa, near-flat responses to 2 kHz, and dimensions below 3 mm — fully compatible with miniaturized chest piece designs. Low power consumption (typically <1 mW active) and direct integration with digital interfaces (I2S, PDM) facilitate seamless coupling with modern SoC processors without discrete AFE components [15,16].

Advanced systems combine MEMS acoustic microphones with MEMS accelerometers or contact vibration sensors in hybrid architectures. Simultaneous multi-modal sensing enables independent component analysis (ICA) for motion artefact rejection — a critical advantage in intensive care units and emergency departments where ambient sound pressure levels routinely exceed 70 dB(A). Multi-element microphone arrays (3–6 capsules) further enable spatial beamforming, selectively enhancing signals from the thoracic surface while suppressing laterally incident noise.

Table 1: Sensor and hardware components of an AI-powered digital stethoscope.

Piezoelectric contact mic	Mechanical → electrical transduction; flat response 20–500 Hz; robust coupling	Cardiac low-frequency sounds (S1/S2)	EKO CORE; Thinklabs One
MEMS condenser mic	SNR ≥65 dB; bandwidth 20–2000 Hz; dimensions <3 mm; low power	Full cardiac + pulmonary spectrum	TI OPA391; Knowles FG-23329
MEMS + accelerometer hybrid	Simultaneous acoustic + vibration sensing; motion artefact rejection	Noisy environments (ICU, ED)	Analog Devices ADXL355
Analog front-end (AFE)	24-bit ADC; PGA; anti-aliasing LPF (fc 2 kHz); 16–22 kHz Fs	Signal fidelity; Nyquist compliance	TI ADS1299; Cirrus CS5368
Edge AI SoC / NPU	INT8 inference; <50 ms latency; <200 mW; ARM Ethos / Hexagon DSP	Real-time on-device classification	Qualcomm APQ8053; NXP i.MX RT
Wireless (BLE 5.2 + Wi-Fi 6)	AES-128 encryption; <20 ms BLE latency; IEEE 802.11ax LAN	Telemonitoring; E-care integration	Nordic nRF5340; TI CC2652

3.2 Analog Front-End and Analog-to-Digital Conversion

The analog front-end (AFE) comprises a low-noise amplifier, a programmable-gain amplifier (PGA), an anti-aliasing low-pass filter, and a high-resolution ADC. A minimum sampling rate of 8 kHz satisfies the Nyquist criterion for cardiac sounds (<4 kHz content), but 16–22.05 kHz is recommended to capture the full lung sound spectrum without aliasing. A 24-bit ADC provides a theoretical dynamic range of approximately 144 dB, eliminating quantization noise as a limiting factor across the physiological amplitude range of body sounds (~60 dB) [16].

Careful PCB layout — including star-ground topology, power supply decoupling (100 nF ceramic + 10 μF electrolytic at each supply pin), and separation of analog and digital ground planes — is mandatory to prevent 50/60 Hz mains hum injection. Shielded cables or fully differential signal paths should be implemented wherever lead lengths exceed 10 cm.

4. Signal Processing Pipeline

4.1 Noise Reduction

Signal acquisition represents only the first step in digital auscultation. Raw acoustic data must undergo several processing stages before clinical interpretation or AI analysis becomes possible. Environmental sounds, friction artefacts, speech contamination, and patient motion can profoundly alter signal quality. Classical techniques — band-pass filtering, adaptive filtering, and wavelet denoising — remain widely deployed because of their low computational cost and predictable behavior [17,18].

The Wiener filter and its spectral subtraction variants achieve effective stationary noise suppression but assume stationarity of the noise field — an assumption routinely violated in clinical environments. Deep learning-based denoisers (e.g., RNNoise, DeepFilterNet) trained on clinical noise corpora handle non-stationary interference more robustly. A two-stage approach — adaptive filtering for stationary components followed by a neural denoiser for residual non-stationary artefacts — reduces downstream AI classification error by approximately 15–20% in reported benchmarks [19,20].

4.2 Feature Extraction

Feature extraction transforms acoustic waveforms into mathematically interpretable representations. Mel-frequency cepstral coefficients (MFCCs), derived from a mel-filterbank applied to the short-time Fourier transform (STFT), compress spectrotemporal information into a compact 13–40 coefficient vector that mimics human auditory perception and remains one of the most widely used input features for conventional classifiers [21,22].

Mel-spectrograms (log-power output of the mel-filterbank without cepstral compression) preserve richer spectrotemporal detail and are preferred as CNN input by treating the spectrogram as a 2-D image. The constant-Q transform (CQT), with geometrically spaced frequency bins, offers superior resolution at low frequencies and is particularly suited to cardiac sound analysis below 200 Hz. Wavelet transforms provide adaptive time-frequency resolution and are effective for detecting transient events such as fine crackles [23].

4.3 Cardiac Cycle Segmentation

Reliable segmentation of S1 and S2 — delineating systole from diastole — is a prerequisite for accurate murmur classification. Classical approaches use the phonocardiogram envelope computed via the Hilbert transform or homomorphic filtering. Deep learning segmentation models — notably bidirectional LSTM and U-Net architectures — achieve sensitivity and specificity >95% for S1/S2 localization on *PhysioNet 2016* benchmark data [24,25].

Synchronized ECG acquisition — feasible via textile dry electrodes integrated into the chest piece — provides an independent timing reference that substantially improves segmentation robustness in patients with irregular rhythms such as atrial fibrillation. The combination of PCG segmentation and ECG gating reduces murmur classification error by approximately 8–12% in preliminary studies.

5. Artificial Intelligence Architectures

5.1 From Handcrafted Features to Deep Learning

Early AI approaches relied on handcrafted acoustic features — MFCCs, zero-crossing rate, spectral centroid — combined with conventional classifiers such as support vector machines (SVMs) or random forests. Although these methods provided proof of concept and remain computationally inexpensive for deployment on low-power microcontrollers, their generalization across devices, clinical settings, and patient populations often remained limited [26].

Deep learning architectures now dominate the field. Convolutional neural networks (CNNs) are particularly effective for spectrogram analysis because they can identify local temporal-frequency patterns associated with murmurs, wheezes, or crackles without explicit feature engineering. VGGNet, ResNet, and EfficientNet architectures pre--trained on ImageNet and fine-tuned on cardiopulmonary sound datasets consistently achieve AUC values of 0.90–0.96 for binary murmur detection. Lightweight 1-D CNNs operating directly on waveform segments reduce inference latency and are attractive for ultra-low-power microcontroller deployments [27,28].

5.2 Sequential and Attention-Based Models

Recurrent neural networks and long short-term memory (LSTM) architectures are well suited for sequential analysis of dynamic cardiopulmonary signals, capturing the rhythmic temporal structure of the cardiac cycle. Bidirectional LSTM models, processing signals in both temporal directions, improve S1/S2 boundary detection by exploiting future context. Gated recurrent unit (GRU) networks offer a computationally lighter alternative with comparable performance on most auscultation benchmarks [29].

Transformer architectures and multimodal foundation models represent the newest frontier. The Audio Spectrogram Transformer (AST) and similar self-attention models applied to mel-spectrograms have reported AUC values up to 0.97 on benchmark datasets, at the cost of substantially higher parameter counts. Future digital stethoscopes may integrate audio, ECG, oxygen saturation, and clinical metadata in unified transformer architectures — functioning as intelligent clinical copilots capable of generating diagnostic suggestions and structured reports [30,31].

5.3 Training Strategies

The quality and diversity of the training corpus are the dominant determinants of model generalization. Major publicly available datasets include *PhysioNet/CinC 2016* (3,126 PCG recordings), the *PASCAL Heart Sound Challenge*, the *ICBHI 2017* respiratory sound database (920 recordings, 6,898 annotated respiratory cycles), and the *SPR-Sound paediatric* database [32,34].

Table 2: AI architectures used in digital stethoscope development.

SVM / Random Forest	Handcrafted MFCC + statistical features	Murmur binary detection (proof of concept)	AUC ~0.82–0.88	Limited generalization
1-D / 2-D CNN	Spectrogram or waveform; ResNet / EfficientNet	Murmur, crackle, wheeze detection	AUC 0.90–0.96	Large labelled dataset needed
LSTM / BiLSTM	Sequential modelling; temporal cardiac cycles	Arrhythmia; S1/S2 segmentation	F1 0.88–0.93	Long training; vanishing gradient
Transformer / AST	Multi-head self-attention; mel-spectrogram	Multi-pathology; COPD + HF combined	AUC up to 0.97	Compute-intensive
Hybrid CNN-LSTM	Spatial + temporal feature fusion	Complex multi-sound classification	F1 ~0.91	Architecture tuning complexity
Federated + SSL	Privacy-preserving distributed training; SimCLR/MAE pretraining	Multi-site deployment; low-resource	Comparable to centralized	Communication & fine-tuning overhead

Class imbalance — abnormal sounds typically constitute only 20–40% of clinical recordings — is addressed through oversampling (SMOTE in feature space), cost-sensitive loss functions, or focal loss. Data augmentation in the time-frequency domain (Spec Augment, pitch shifting, time stretching, synthetic noise injection) improves generalization by 3–8 percentage points in F1-score. Self-supervised learning (SimCLR, masked audio auto encoder) enables pre-training on large unlabeled corpora, markedly reducing the volume of expert-annotated data required for fine-tuning [35,36].

Federated learning — in which model weights rather than raw audio are shared across hospital nodes — enables privacy-preserving multi-center training while satisfying GDPR constraints. Early implementations for cardiac sound classification demonstrate performance comparable to centralized training when ten or more federated sites participate, and represent the preferred architecture for large-scale ASAP-type deployments. [37,38].

6. Clinical Applications

6.1 Cardiovascular Medicine

Cardiovascular medicine represents the most mature domain for AI-enhanced digital auscultation. Deep learning systems have demonstrated promising performance for the detection of valvular heart disease, systolic dysfunction, and heart failure. Aortic stenosis, the most prevalent valvular disease in adults aged over 65 years, produces a characteristic ejection systolic murmur detectable with high sensitivity by trained AI models; large-scale opportunistic screening programs in general practice have been proposed to reduce the estimated 1.5–2% prevalence of undiagnosed severe aortic stenosis in adults over 75 year [39,40].

Pulmonary crackles are an early, sensitive marker of rising pulmonary capillary wedge pressure in decompensating heart failure. Automated daily crackle surveillance in ambulatory patients with reduced ejection fraction — transmitted via the *E-care* platform to a heart failure nurse — has shown potential for detecting decompensation 3–5 days before clinical symptom onset, enabling preemptive diuretic adjustment [41].

6.2 Pulmonary Medicine

AI systems can identify wheezes, crackles, rhonchi, and other abnormal respiratory sounds associated with asthma, COPD, pneumonia, or interstitial lung disease. During the COVID-19 pandemic, remote auscultation and telemedicine applications attracted considerable interest because they reduced direct clinician exposure while preserving clinical monitoring capability, accelerating the validation of connected digital stethoscopes in real-world settings [42,43].

Table 3: Selected clinical validation studies of AI-enhanced digital stethoscopes.

Clifford et al. 2016	PhysioNet/CinC Challenge	Murmur (generic)	n = 3,126 PCG	AUC 0.86 (winner)	Heterogeneous labels
Chorba et al. 2021	DL via digital stethoscope	Systolic dysfunction (LVEF <40%)	n = 2,056	AUC 0.93	Retrospective
Gardezi et al. 2021	CNN; multi-label	Valvular heart disease	n = 1,683	Sen 87%, Spe 91%	Specialist annotation
Kevat et al. 2020	Digital stethoscope + AI	Paediatric breath sounds	n = 345 children	AUC 0.89	Single centre
Andrès et al. 2024	eStetho (ASAP / HUS)	Internal medicine (full spectrum)	n = 857	Kappa ~0.87; ~94% concordance	PRI HUS No. 4179
Smail et al. 2023	AI-enhanced stethoscope review	Multi-domain (cardiac + pulm.)	Systematic review	Narrative synthesis	Heterogeneous designs

6.3 Pediatrics, Primary Care, and Resource-Limited Settings

Congenital heart disease affects approximately 8 per 1,000 live births; up to 30% of hemodynamically significant defects remain undetected at hospital discharge in lowincome countries without echocardiographic access. AI stethoscopes trained on pediatric normative data and CHD-specific murmur patterns represent a transformative screening tool for primary care workers in these settings, enabling referral triage without specialist availability [44,45].

Continuous wireless auscultation in hospitalized patients — facilitated by wearable chest piece designs — may contribute to early detection of clinical deterioration, complementing pulse oximetry and telemetric ECG in a multiparameter early-warning system. The democratization of low-cost AI-enhanced digital stethoscopes carries significant global health implications, potentially improving access to diagnostic expertise in resource-limited regions through remote specialist consultation.

7. Engineering Challenges

7.1 Energy Autonomy and Embedded Computation

Battery autonomy remains a central constraint for portable and continuous monitoring applications. Neural processing units embedded in modern SoCs support INT8 quantized inference at below 200 mW, enabling all-day operation from a 500 mAh lithium-polymer cell. Model compression techniques — pruning, knowledge distillation, and quantization-aware training — reduce parameter counts by 50–80% with less than 2% accuracy loss, enabling full AI inference on Cortex-M class microcontrollers [46].

7.2 Connectivity and Cybersecurity

Bluetooth Low Energy 5.2 ensures encrypted audio streaming to a companion smartphone application with below 20 ms latency; IEEE 802.11ax (Wi-Fi 6) supports hospital-grade LAN integration. As medical devices exchange sensitive patient information through digital infrastructures, cybersecurity and data protection become increasingly critical design requirements. AES-128 encryption at the transport layer, certificate-based mutual authentication, and over-the-air (OTA) firmware update mechanisms with cryptographic signature verification are the minimum acceptable security posture for a CE-marked or FDA-cleared device [47].

7.3 Human-Centered Design

Human-centered design plays a decisive role in clinical adoption. Physicians require ergonomic devices that integrate naturally into clinical workflows without increasing cognitive burden. Excessively complex interfaces, mandatory connectivity steps before use, or lengthy boot sequences will reduce adoption regardless of algorithmic performance. User studies involving both nursing staff and senior physicians at the prototype stage — using standardized usability frameworks such as the System Usability Scale (SUS) — should be considered mandatory prior to clinical validation. Engineers must balance technological sophistication against simplicity and intuitiveness, ensuring that AI outputs are presented as actionable clinical probabilities rather than raw logits or opaque scores.

8. Clinical Validation and Regulatory Pathways

8.1 Performance Metrics and Reporting Standards

Clinical validation must report, at minimum, sensitivity, specificity, positive and negative predictive values, AUC of the ROC curve, and Cohen's kappa for agreement with expert reference. Calibration metrics — reliability diagrams and expected calibration error (ECE) — should complement discrimination metrics, as well-calibrated confidence scores are essential for safe clinical decision support. Systematic reporting of explainability analyses (Grad-CAM, SHAP, LRP) should be considered mandatory, in alignment with the STARD-AI extension of the STARD reporting guideline [48,49].

8.2 Regulatory Framework

AI-enhanced medical devices are subject to increasingly rigorous regulatory frameworks. In the European Union, digital stethoscopes are classified as Class IIa medical devices under MDR 2017/745, requiring Notified Body conformity assessment and application of the EN ISO 13485 quality management system. The *EU AI Act* (2024) further imposes obligations specific to high-risk AI systems used in patient diagnosis, including mandatory human oversight, transparency, robustness testing, and registration in the EU AI database [50].

In the United States, the FDA regulates AI-enabled stethoscopes under the De Novo or 510(k) pathways. The FDA's 2021 action plan for AI/ML-based software as a medical device (SaMD) introduced predetermined change control plans (PCCPs) to facilitate iterative model updates without full regulatory re-submission — a critical enabler for continuously learning systems.

8.3 Ethical Considerations

Ethical concerns are as consequential as regulatory ones. Algorithmic bias may reduce diagnostic performance in under-represented populations — those with higher BMI, non-European thoracic morphology, or rare phenotypes absent from training corpora. Automation bias represents a distinct and equally serious risk: physicians may over-rely on algorithmic outputs and abandon active clinical reasoning. The objective should never be to replace clinical judgement but to augment diagnostic capabilities while preserving physician autonomy and accountability [3,48].

Table 4: Regulatory pathways for AI-enhanced digital stethoscopes.

EU (MDR 2017/745)	Class IIa medical device	Notified Body conformity; EN ISO 13485 QMS	Post-market clinical follow-up; vigilance reporting
EU AI Act 2024	High-risk AI system (Annex III)	Human oversight; transparency; robustness testing	Mandatory registration in EU AI database
USA (FDA SaMD)	De Novo or 510(k) pathway	Predetermined Change Control Plan (PCCP)	Real-world performance monitoring; updates require notification
France (HDS/ANSM)	Hébergeur de Données de Santé certification	Data sovereignty; pseudonymisation	Interoperability with DMP (Dossier Médical Partagé)

9. Telemedicine Integration: The E-Care Platform

The E-care platform, developed at the *Hôpitaux Universitaires de Strasbourg* (France), provides a cloud-based telemonitoring infrastructure designed to aggregate data from multiple connected medical devices, including the ASAP digital stethoscope. Data are transmitted over AES-128 encrypted BLE channels to a gateway application, then forwarded via HTTPS to a pseudo anonymized server compliant with the French HDS (*Hébergeur de Données de Santé*) certification framework [4,5].

The platform aggregates auscultatory findings with pulse oximetry, blood pressure, body weight, and ECG streams, presenting an integrated dashboard to the supervising internist. Configurable alert thresholds trigger asynchronous push notifications for high-probability pathological auscultation events, enabling timely clinical review without requiring real-time physician availability. This architecture supports both planned teleconsultations and unscheduled remote assessments — an essential feature for the management of frail, elderly, or geographically isolated patients [5].

Integration with the hospital information system via HL7 FHIR R4 ensures that AI-generated auscultatory reports are stored in the patient's electronic health record and are fully auditable. Semantic interoperability is achieved through mapping to SNOMED CT concepts for cardiac and respiratory findings, facilitating secondary data use for research and epidemiological surveillance. Future digital stethoscopes may therefore function not only as diagnostic aids but as persistent longitudinal biosensors contributing to population health monitoring.

10. Future Perspectives

Future digital stethoscopes will probably evolve into multimodal physiological monitoring platforms integrating electrocardiography, oxygen saturation, respiratory mechanics, body temperature, and wearable biosensors into a single ergonomic chest piece. Continuous home monitoring enabled by such platforms may transform the management of chronic cardiovascular and pulmonary diseases, shifting the locus of care from episodic hospital encounters to continuous community-based surveillance [1,3].

Self-supervised learning and federated learning will improve algorithmic robustness while preserving patient privacy at scale. Large-scale foundation models pre--trained on multi-million recording corpora — analogous to GPT-class language models — could be fine-tuned for any auscultatory task with minimal labelled data, dramatically accelerating the development cycle for new clinical indications. Explainable AI will be indispensable for clinical adoption: physicians require transparent systems that justify their predictions through saliency maps, attention weights, or natural language rationales rather than opaque probability scores [37,49].

The fusion of auscultatory AI with point-of-care ultrasound algorithms, natriuretic peptide point-of-care testing, and real-time respiratory function monitoring may eventually enable a comprehensive, portable, non-invasive cardiopulmonary assessment platform accessible to any healthcare worker worldwide — realizing the full democratization potential of AI-enhanced clinical instrumentation.

11. Conclusion

The AI-enhanced digital stethoscope illustrates the convergence of medicine, engineering, acoustics, and data science in one of the most recognizable instruments in clinical practice. These devices are transforming auscultation from a subjective clinical art into a quantitative, reproducible, and remotely transmissible diagnostic modality, without sacrificing the gestural intimacy of the clinical encounter.

The future success of digital auscultation will depend not only on algorithmic performance but equally on interoperability, explainability, regulatory validation, equitable training data, and seamless clinical integration. Collaboration between physicians and engineers remains essential and irreplaceable: physicians provide clinical relevance, phenotypic annotation, and diagnostic reasoning; engineers contribute expertise in signal processing, embedded systems, software architecture, and machine learning. Neither discipline alone is sufficient.

Rather than replacing bedside medicine, AI-enhanced digital stethoscopes can revitalize the clinical examination by merging traditional medical observation with advanced computational intelligence. The ASAP/E-care experience in Strasbourg demonstrates the feasibility of high-performance, clinically integrated AI auscultation within an academic medical center setting — a model applicable to centers worldwide seeking to elevate diagnostic quality at the point of care.

Conflict of Interest Statement

The authors declare no conflicts of interest relevant to this article. N. Lorenzo-Villalba serves on the Scientific Committee of La Revue de Médecine Interne (COPE declaration). E. Andrès serves as Chef Editor of Journal of Clinical Medicine (COPE declaration).

Funding

No specific funding was received for this review.

Author Contributions

E.A.: concept, literature review (using Pubmed, Google Scholar, ResearchGate), manuscript drafting, final approval. N.L.-V.: literature review, critical revision, co-authorship. A.H.E.H.: IA contribution, critical revision. All authors approved the final version of the manuscript. Generative artificial intelligence tools, including ChatGPT and Claude AI, were used during the preparation of this manuscript for language editing, stylistic improvement, text structuring, and literature synthesis assistance. All AI-assisted content was critically reviewed, revised, and validated by the authors, who take full responsibility for the accuracy, integrity, and originality of the final manuscript. AI tools were not used for data generation, data analysis, or independent scientific decision-making.

References

1. Topol EJ (2019) High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 25: 44-56.

2. Rajpurkar P, Chen E, Banerjee O, Topol EJ (2022) AI in health and medicine. Nat Med. 28: 31-8.

3. Beam AL, Kohane IS (2018) Big data and machine learning in health care. JAMA. 319: 1317-8.

4. Andrès E, Talha S, Hajjam M, et al. (2022) Current Perspective on the Use of Digital Health and Telemedicine in the Management of Heart Failure. J Clin Med. 11: 5008.

5. Andrès E, Talha S, Hajjam El Hassani A, et al. (2023) Telemonitoring in Internal Medicine: Current Perspective. J Clin Med. 12: 2557.

6. Tavel ME (2006) Cardiac auscultation: a glorious past — and it does have a future. Circulation. 113: 1255-9.

7. Iung B, Vahanian A (2011) Epidemiology of valvular heart disease in the adult. Nat Rev Cardiol. 8: 162-72.

8. Etchells E, Bell C, Robb K (1997) Does this patient have an abnormal systolic murmur? JAMA. 277: 564-71.

9. Pasterkamp H, Kraman SS, Wodicka GR (1997) Respiratory sounds. Advances beyond the stethoscope. Am J Respir Crit Care Med. 156: 974-87.

10. Leng S, Tan RS, Allen JC, et al. (2015) The electronic stethoscope. Biomed Eng Online. 14: 66.

11. Piirilä P, Sovijärvi AR (1995) Crackles: recording, analysis and clinical significance. Eur Respir J. 8: 2139-48.

12. Gurung A, Scrafford CG, Tielsch JM, Levine OS, Checkley W (2011) Computerized lung sound analysis as diagnostic aid for the detection of abnormal lung sounds: a systematic review and meta-analysis. Respir Med. 105: 1396-403.

13. Loudon R, Murphy RL Jr (1984) Lung sounds. Am Rev Respir Dis. 130: 663-73.

14. Watrous RL (2006) Computer-aided auscultation of the heart: from anatomy and physiology to diagnostic decision support. IEEE Eng Med Biol Mag. 25: 16-26.

15. Leng S, Tan RS, Allen JC, et al. (2015) The electronic stethoscope. Biomed Eng Online. 14: 66.

16. Sovijärvi AR, Vanderschoot J, Earis JE (2000) Standardization of computerized respiratory sound analysis. Eur Respir Rev. 10: 585-649.

17. Widrow B, Stearns SD (1985) Adaptive Signal Processing. Englewood Cliffs: Prentice-Hall.

18. Pramono RXA, Imtiaz SA, Rodriguez-Villegas E (2017) Automatic adventitious respiratory sound analysis: a systematic review. PLoS One. 12: e0177926.

19. Lim JS, Oppenheim AV (1979) Enhancement and bandwidth compression of noisy speech. Proc IEEE. 67: 1586-604.

20. Goodfellow I, Bengio Y, Courville A (2016) Deep Learning. Cambridge: MIT Press.

21. Davis S, Mermelstein P (1980) Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. IEEE Trans Acoust. 28: 357-66.

22. Chen W, Sun Q, Chen X, et al. (2021) Deep Learning Methods for Heart Sounds Classification: A Systematic Review. Entropy. 23: 667.

23. Demir F, Sengur A, Bajaj V (2020) Convolutional neural networks based efficient approach for classification of lung diseases. Health Inf Sci Syst. 8: 4.

24. Springer DB, Tarassenko L, Clifford GD (2016) Logistic Regression-HSMM-Based Heart Sound Segmentation. IEEE Trans Biomed Eng. 63: 822-32.

25. Roth GA, Mensah GA, Johnson CO, et al. (2020) Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019. J Am Coll Cardiol. 76: 2982-3021.

26. Potes C, Parvaneh S, Rahman A, Conroy B (2016) Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds. Comput Cardiol. 43:

621-4.

27. Chorba JS, Shapiro AM, Le L, et al. (2021) Deep learning algorithm for automated cardiac murmur detection via a digital stethoscope platform. J Am Heart Assoc. 10: e019905.

28. Clifford GD, Liu C, Moody B, et al. (2016) Classification of normal/abnormal heart sound recordings: The PhysioNet/Computing in Cardiology Challenge. Comput Cardiol. 43: 609-12.

29. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput. 9: 1735-80.

30. Gong Y, Chung YA, Glass J (2021) AST: Audio Spectrogram Transformer. arXiv:2104.01778.

31. Vaswani A, Shazeer N, Parmar N, et al. (2017) Attention Is All You Need. Adv Neural Inf Process Syst. 30.

32. Rocha BM, Filos D, Mendes L, et al. (2018) A respiratory sound database for the development of automated classification. Precision Medicine Powered by pHealth. Springer; 33-7.

33. Zhang J, Tao J, Zhang W, et al. (2022) SPRSound: Open-Source SJTU Paediatric Respiratory Sound Database. IEEE Trans Biomed Circuits Syst. 16: 867-81.

34. Kevat A, Kalirajah A, Roseby R (2020) Artificial intelligence accuracy in detecting pathological breath sounds in children using digital stethoscopes. Respir Res. 21: 156.

35. Park DS, Chan W, Zhang Y, et al. (2019) SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition. Proc Interspeech. 2613-7.

36. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature. 521: 436-44.

37. Rieke N, Hancox J, Li W, et al. (2020) The future of digital health with federated learning. NPJ Digit Med. 3: 119.

38. McMahan B, Moore E, Ramage D, Hampson S, Agüera y Arcas B (2017) Communication-Efficient Learning of Deep Networks from Decentralized Data. Proc AISTATS. 1273-82.

39. Gardezi SKM, Myerson SG, Chambers J, et al. (2018) Cardiac auscultation poorly predicts the presence of valvular heart disease in asymptomatic patients with cardiovascular risk factors. Heart. 104: 1832-5.

40. Iung B, Baron G, Butchart EG, et al. (2003) A prospective survey of patients with valvular heart disease in Europe: The Euro Heart Survey on Valvular Heart Disease. Eur Heart J. 24: 1231-43.

41. Maisel AS, Duran JM, Wettersten N (2018) Natriuretic Peptides in Heart Failure. Heart Fail Clin. 14: 13-25.

42. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries, 1990-2019. Lancet. 396: 1204-22.

43. Andrès E, Hajjam M, Talha S, et al. (2022) Use of Connected Devices and Digital Medicine in Management of Patients with Type 2 Diabetes. J Clin Med. 11: 4351.

44. Norrish G, Ding T, Field E, et al. (2019) Development of a novel risk prediction model for sudden cardiac death in childhood hypertrophic cardiomyopathy (HCM Risk-Kids). JAMA Cardiol. 4: 918-27.

45. Smail H, Andrès E, Talha S, Hajjam El Hassani A (2023) Artificial Intelligence-Based Stethoscope for the Diagnosis of Heart Disease: A Narrative Review. J Clin Med. 12: 476.

46. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput. 9: 1735-80.

47. European Parliament. Regulation (EU) 2017/745 on Medical Devices (MDR). Off J Eur Union.

48. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2020) Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. Int J Comput Vis. 128: 336-59.

49. Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 30.

50. European Parliament. Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (AI Act).

Off J Eur Union. 2024.

Submit your manuscript to a JScholar journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Better discount for your subsequent articles

Submit your manuscript at
http://www.jscholaronline.org/submit-manuscript.php