Audio-based digital biomarkers in diagnosing and managing respiratory diseases: a systematic review and bibliometric analysis

Vivianne Landry [1,2], Jessica Matschek [2,3], Roger Pang[2,3], Meghana Munipalle[4], Kenneth Tan[3], Jill Boruff[5] and Nicole Y.K. Li-Jessen[2,4,6,7]

[1]Faculty of Medicine, University of Montreal, Montreal, QC, Canada. [2]School of Communication Sciences and Disorders, Faculty of Medicine and Health Sciences, McGill University, Montreal, QC, Canada. [3]Faculty of Medicine and Health Sciences, McGill University, Montreal, QC, Canada. [4]Department of Biomedical Engineering, Faculty of Medicine and Health Sciences, McGill University, Montreal, QC, Canada. [5]Schulich Library of Physical Sciences, Life Sciences, and Engineering, McGill University, Montreal, QC, Canada. [6]Department of Otolaryngology-Head and Neck Surgery, Faculty of Medicine and Health Sciences, McGill University, Montreal, QC, Canada. [7]Translational Research in Respiratory Diseases Program, The Research Institute of the McGill University Health Centre, Montreal, QC, Canada.

Corresponding author: Nicole Y.K. Li-Jessen (nicole.li@mcgill.ca)

Abstract

Advances in wearable sensors and artificial intelligence have greatly enhanced the potential of digitised audio biomarkers for disease diagnostics and monitoring. In respiratory care, evidence supporting their clinical use remains fragmented and inconclusive. This study aimed to assess the current research landscape of digital audio biomarkers in respiratory medicine through a bibliometric analysis and systematic review (PROSPERO CRD 42022336730). MEDLINE, Embase, Cochrane Library and CINAHL were searched for references indexed up to 9 April 2024. Eligible studies evaluated the accuracy of sound analysis for diagnosing and managing obstructive (asthma and COPD) or infectious respiratory diseases, excluding COVID-19. A narrative synthesis was conducted, and the QUADAS-2 tool was used to assess study quality and risk of bias. Of 14 180 studies, 81 were included. Bibliometric analysis identified fundamental (*e.g.* "diagnostic accuracy"+"machine learning") and emerging (*e.g.* "developing countries") themes. Despite methodological heterogeneity, audio biomarkers generally achieved moderate (60–79%) to high (80–100%) accuracies. 80% of studies (eight out of ten) reported high sensitivities and specificities for asthma diagnosis, 78% (seven out of nine) reported high sensitivities and 56% (five out of nine) reported high specificities for COPD, and 64% (seven out of eleven) reported high sensitivity or specificity values for pneumonia diagnosis. Breathing and coughing were the most common biomarkers, with artificial neural networks being the most common analysis technique. Future research on audio biomarkers should focus on testing their validity in clinically diverse populations and resolving algorithmic bias. If successful, digital audio biomarkers hold promise for complementing existing clinical tools in enabling more accessible applications in telemedicine, communicable disease monitoring, and chronic condition management.

Background

Respiratory diseases constitute one of the leading causes of morbidity and mortality worldwide. Chronic obstructive pulmonary disease (COPD), asthma, tuberculosis and lower respiratory tract infections stand out as some of the most common causes of severe illness worldwide, with nearly 7.5 million associated deaths every year [1]. Early diagnosis and intervention are key to mitigating the adverse impacts of these conditions on both individuals and societies [2]. Stethoscopic lung auscultation is a widely recognised

diagnostic tool for detecting pulmonary symptoms given its simplicity, safety, portability, cost-effectiveness and wide accessibility [3–5]. However, lung auscultation is highly subjective and its accuracy relies on the provider's experience and perceptual abilities [5–8]. Incorrect or delayed diagnoses have been reported even among experienced healthcare professionals [5, 7, 8].

Complementary diagnostic modalities including but not limited to medical imaging, spirometry, bronchoscopy, blood work and sputum cultures are available for confirming a diagnosis initially suggested by anamnesis and physical examination. These supplementary methods can sometimes be costly, poorly accessible in rural regions with limited resources, and may pose challenges in terms of patient tolerance and acceptability [9, 10]. In addition, for communicable respiratory diseases such as COVID-19, new technologies that support contact-free remote diagnosis are in high demand [11].

Recent advances in wearable devices and artificial intelligence (AI) hold potential for harnessing clinically relevant sounds in digital health technologies [12–16]. An array of body sounds such as speech, heartbeats, breathing and coughs has been proposed for respiratory disease diagnosis and management. These sounds are acquired through audio-acoustic microphones (*e.g.* embedded in mobile devices) or mechano-acoustic sensors (*e.g.* wearable accelerometers) in both clinical and free-living settings [12–14, 17]. Owing to the complexity of physiological sounds, machine learning (ML)-based or AI-based algorithms have greatly automated the analytic pipeline and made clinical evaluation more efficient than human inspection alone [13, 14]. Classifiers of sound signals in conjunction with a broad spectrum of ML algorithms have demonstrated promising results in the detection of respiratory disorders [13, 14, 18]. Common applications include automated counts of cough sounds, discrimination of wheezes and crackles from normal lung sounds, and detection of sleep apnoea through snore signals [19–22].

Previous research has mostly focused on the automatic detection of abnormal lung sound characteristics, such as wheezes or coughs, from the collected audio signals. Very few studies have tested the use of abnormal lung sounds in diagnosis, classification, monitoring and symptom management of respiratory conditions.

In this review, we sought to provide and synthesise the latest evidence on the research landscape, reliability, and validity of digital audio-based biomarkers in respiratory medicine through a bibliometric analysis and systematic review. Specific study objectives were to 1) evaluate research trends and outputs on the digital health technology of audio-based biomarkers in respiratory medicine; and 2) evaluate the performance of digital audio biomarkers, their acquisition technology, and data analytical methods in diagnosing and managing prevalent respiratory disorders.

Methods
This review was registered on the International Prospective Register of Systematic Reviews (PROSPERO: CRD 42022336730). The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) framework was used to guide the reporting of this review [23]. A systematic review and bibliometric analysis were performed to identify relevant studies reporting on the accuracy of sound analysis in diagnosing and managing obstructive diseases (*i.e.* asthma and COPD) and infectious respiratory diseases (*e.g.* pneumonia, bronchitis, upper respiratory tract infection). These respiratory conditions were selected due to the importance of their global burden, representing respectively the third and fourth leading causes of death globally, and due to the associated potential clinical utility of acoustic-based diagnostic methods, particularly in resource-limited settings [1].

Search strategy
A health sciences librarian created a search strategy in Ovid MEDLINE that was peer-reviewed by a second librarian using the Peer Review of Electronic Search Strategies guidelines [24]. The librarian then adapted the searches for Ovid Embase, CINAHL (EBSCO) and the Cochrane Library (Wiley). A filter for diagnostic accuracy was built using the Scottish Intercollegiate Guidelines Network diagnostic filter, the GEERSING *et al*. [25] additional line, the HAYNES *et al*. [26] balanced diagnosis filter, and terms from textual analysis of selected titles and abstracts using the Text Analyzer website (www.online-utility.org/text/analyzer.jsp). The complete search strategies can be found in the supplementary material and in the following institutional repository: https://doi.org/10.5683/SP3/GMHZ1Y.

The search was first conducted on 22 August 2022 and refreshed on 9 April 2024. This review included all eligible studies available in the databases up to the latest search date of 9 April 2024. References identified by manual search, expert recommendations or within the reference lists of included studies were also considered.

Screening and eligibility assessment of references

The assessment of references for eligibility followed a two-step procedure. First, titles and abstracts of studies obtained from the search strategy were screened for relevancy by two independent researchers. Subsequently, all pertinent records were assessed by full-text reading and their alignment with the eligibility criteria was examined. Disagreements were resolved through mediation by the senior researcher in the team.

Inclusion criteria

Studies were included if they evaluated the accuracy of any physiological or speech sound analysis (*e.g.* breathing, coughing) for the diagnosis or management (including evaluation of severity, disease control, disease exacerbations and response to treatment) of obstructive diseases (asthma and COPD) and infectious respiratory diseases (excluding COVID-19) in human subjects. This analysis could be conducted manually or by automated means. Studies were required to report at least one accuracy measure, including sensitivity, specificity or area under the receiver operating characteristic curve (AUC). Only articles written in English or French were included due to limited resources for the translation of studies.

Exclusion criteria

Non-empirical studies (*i.e.* editorials, commentaries, reviews) and non-peer-reviewed papers (*e.g.* conference abstracts and conference papers) were excluded from this review. Studies conducted in animal populations were also excluded. Studies focusing on COVID-19 were omitted from this review due to the rapidly evolving diagnostic landscape of this disease. The patient, intervention, comparison, outcome (PICO) table that guided the development of the search strategy and reflects the eligibility criteria has been included as supplementary table S1 [23, 27, 28].

Data extraction

Data extracted from the studies included the following: study country, sample size, patient characteristics, gold standard diagnostic methods used for comparison, sound recorded, device used for audio recording, analytical method, classification features and accuracy values.

Quality assessment and risk of bias

Two independent researchers conducted an evaluation of each included study's methodological quality and risk of bias using the QUADAS-2 tool [29]. Of note, these assessments did not influence the eligibility for inclusion in this review. Applicability of the selection criteria, index and reference tests was reviewed in all studies. Risk of bias was assessed with consideration of four main aspects: patient selection, reference test, index test, and flow and timing of the study.

Bibliometric analysis

The open-source R package *bibliometrix* and its associated web app, *biblioshiny*, were used to conduct bibliometric analysis on the final included studies (n=81) [30]. To ensure that all relevant metadata were included, 80 studies were exported in the BibTex (.bib) file format from the SCOPUS database, and one entry unavailable in SCOPUS was manually added in the same format. All exported studies had the following metadata available: abstract, affiliation, author(s), digital object identifier (DOI), document type, language, publication year, title and total citations. 94% of the studies also included keywords-plus (index terms automatically generated from the references) and corresponding authors. 86% of the studies included author keywords.

A tree map-style visualisation, using the pyBibX bibliometric analysis package in Python, was used to obtain insights into the quality of the journals in which digital audio biomarker research has been published [31]. The top ten most frequently published-in journals for the collection of studies were presented alongside their 5-year journal impact factor (JIF) and CiteScore (CS), reported in the 2023 edition of Journal Citation Reports [32]. A 5-year JIF is the average number of citations received in 2023 by articles and reviews published in the past 5 years in a given journal. This metric is calculated by dividing the total number of citations (2023) received for articles published in the previous 5 years (2018–2022) by the total number of articles published between 2018 and 2022 (5 years). CS is a metric provided using data from SCOPUS and is calculated similarly to the 5-year JIF from Clarivate, but instead uses a 3-year window and includes all items (articles, reviews, conference proceedings, editorials, *etc.*).

Keywords-plus metadata were used to generate a thematic map and trend topics. To provide a clearer view of research topics, a short list of stopwords (*i.e.* excluded keywords) was entered with common but trivial phrases such as human(s), article, *etc*. The bivariate thematic map places clusters along two axes: cluster centrality (*i.e.* relevance) and cluster density (*i.e.* development). Clusters are divided into four groups based

on centrality and density [33, 34]: 1) emerging or declining themes (low density and centrality); 2) niche themes (low centrality and high density; specialised areas of the research field); 3) basic themes (low density and high centrality; general and fundamental themes); and 4) motor themes (high density and centrality; important and well-developed themes).
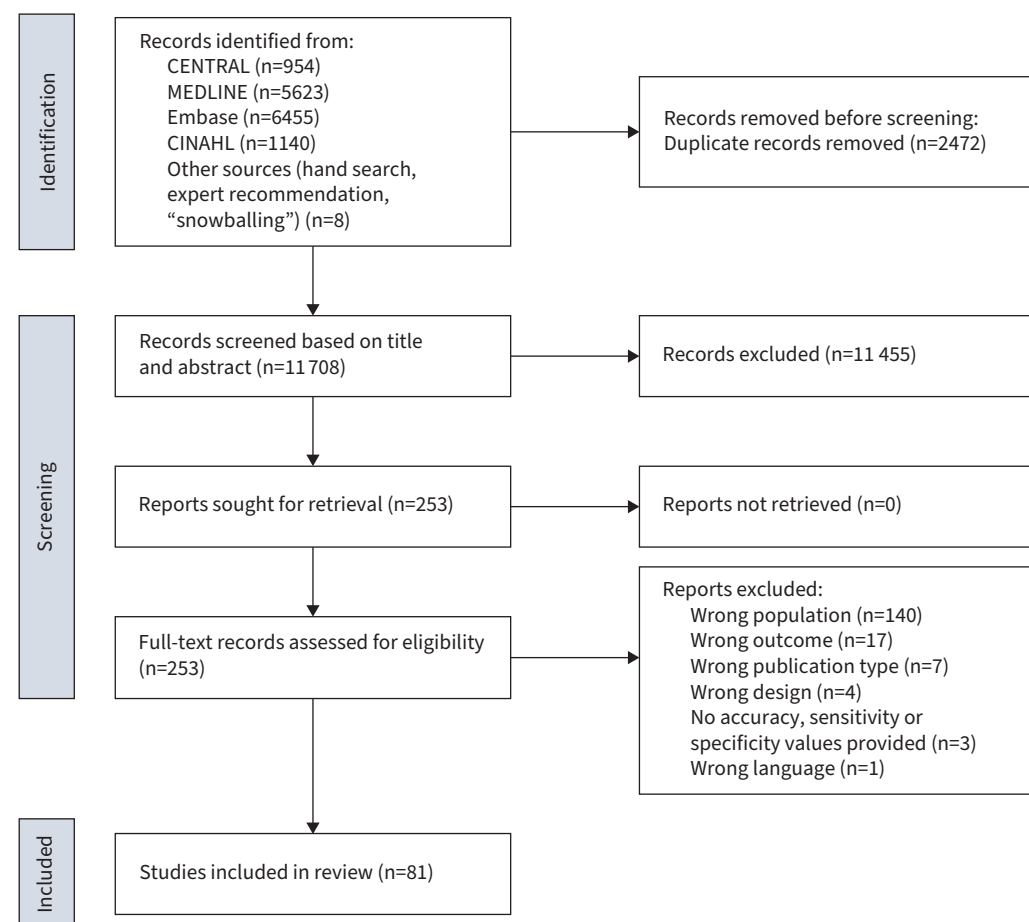
Results

Search and screening results

The search strategy produced a total of 14 172 results, which were subsequently imported into the EndNote citation manager software (version X9; the EndNote Team, Philadelphia, PA, USA). After eliminating 2472 duplicates through EndNote's automation, eight additional records were identified from alternative sources such as reference lists of included studies, manual search and expert recommendations. Following this deduplication process, a total of 11 708 studies were imported into the Rayyan systematic review software for screening [35]. At the end of this process, 253 records were evaluated for eligibility by full-text assessment. Out of these, 81 studies met the criteria for final inclusion (figure 1).

General study characteristics

A total of 81 final studies were included from 27 different countries [36–116]. Studies were mostly from Australia (n=10), India (n=9), the USA (n=8), China (n=7) and Spain (n=5). The remaining studies were from Pakistan (n=4), Turkey (n=4), Japan (n=4), Poland (n=3), Singapore (n=3), Finland (n=2), Indonesia (n=2), Israel (n=2), Jordan (n=2), South Korea (n=2), South Africa (n=2), the UK (n=2), Austria (n=1), Bangladesh (n=1), Ethiopia (n=1), Iraq (n=1), Nepal (n=1), the Netherlands (n=1), Russia (n=1), Switzerland (n=1), Taiwan (n=1) and Vietnam (n=1).

FIGURE 1 Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow chart of search results.

A total of 19 studies were on asthma [36–54], 16 studies were on COPD [55–70] and 19 studies were on infectious respiratory diseases [71–89]. There were 27 studies focusing on conditions that included obstructive and infectious diseases but not limited exclusively to a single one of these categories [90–116]. A total of 20 studies were on paediatric patients [37, 39, 40, 42, 48, 49, 51, 71, 74, 80–82, 84, 86–88, 91, 92, 97, 107], while 21 studies focused on adult patient populations [43–47, 50, 54, 57–60, 64, 65, 73, 75, 85, 89, 95, 96, 100, 105], four included both paediatric and adult participants [38, 52, 53, 93] and 36 did not specify sample age [36, 41, 55, 56, 61–63, 66–70, 72, 76–79, 83, 90, 94, 98, 99, 101–104, 106, 108–116]. For sound analysis, ML techniques were deployed in most studies (n=70) [36–38, 41–47, 51–59, 61–81, 83–86, 88–91, 93, 94, 96–99, 101–116] with a relatively small number of studies using conventional statistical analysis (n=11) [39, 40, 48–50, 60, 82, 87, 92, 95, 100] and non-automated analysis (n=5) [49, 60, 82, 92, 95]. Out of 81 studies, less than 10% (n=8) collected the audio data at home or in remote settings [48, 49, 53, 58, 59, 62, 69, 95]. About 32% (n=26) relied on publicly available data sources such as the International Conference on Biomedical Health Informatics dataset (ICBHI 2017) [117]. Most studies (approximately 60%, n=50) reported their data collection in clinical settings.

Research trends and publication patterns

Based on the bibliometric analysis, both the number of published studies and the average number of citations on digital audio biomarkers in respiratory diseases have notably increased within the last 10 years (figure 2). The top ten journals in which most publications were seen had JIFs between 1.9 and 7.1 and a CS between 4.0 and 16.5 (figure 3a). The journals with the highest number of studies published were *Biomedical Signal Processing and Control* (JIF: 4.9; CS: 9.8) followed by *Sensors* (JIF: 3.7; CS: 7.3). A total of 15 papers were published in these two journals.

A thematic map analysis revealed specific themes in the body of audio biomarker research (figure 3b). The map distinguished between emerging, niche and fundamental research areas and topics. A major cluster in the motor themes category contained the keywords "diagnostic accuracy" and "diagnosis" along with "deep learning" and "machine learning", reflecting the utility of well-developed AI algorithms in improving the diagnostic accuracy of audio biomarker methods. Low density, low centrality clusters contained keywords such as "developing countries", "digital stethoscope" and "e-learning", identifying strong and necessary avenues for further research in equitable digital health beyond the use of ML algorithms. An analysis of trend topics demonstrated the relative recency of deep learning and ML models in the context of audio biomarker analysis (supplementary figure S1).

Risk of bias appraisal

The risk of bias and applicability concerns were evaluated using the QUADAS-2 tool. The quality assessment revealed heterogeneity across studies (figure 4 and supplementary table S2). In this context, risk of bias refers to the potential that study-related factors (*e.g.* design or reporting) may compromise the validity of the findings. Applicability describes the extent to which the study's characteristics align with the review question and its clinical context. Regarding the risk of bias, the "patient selection", "reference standard" and "flow and timing" domains were more often evaluated as "unclear" or "high" (90%, 48% and 46%, respectively), in comparison with the "index test" domain (21%). Overall, 80% of studies
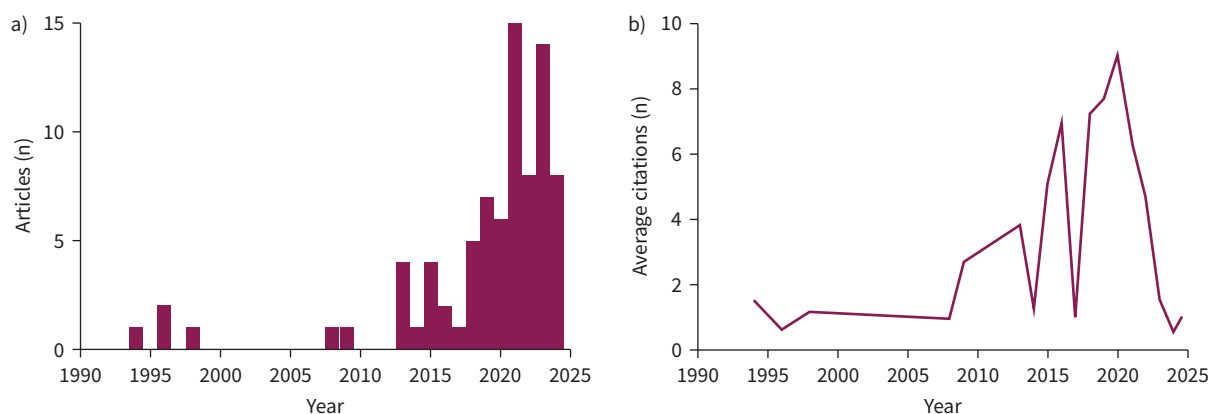

FIGURE 2 Publication trends in audio-based digital biomarkers for respiratory diagnoses. a) Number of studies per year (1994–2024). b) Average number of citations per year.
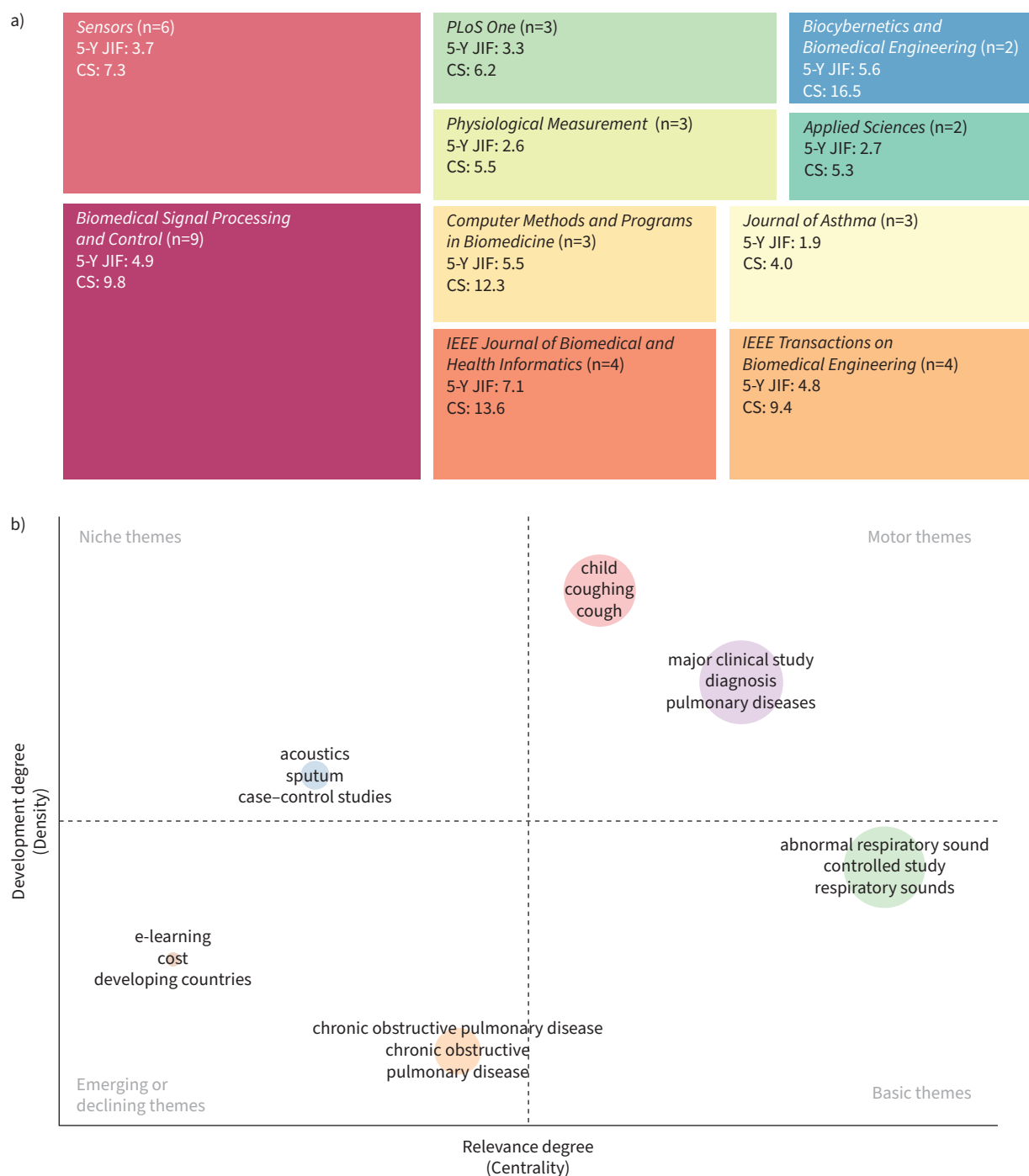
FIGURE 3 Journal and thematic trends in audio-based digital biomarkers for respiratory diagnoses. a) Tree map visualisation of the top ten journals in which the included studies were most frequently published. The 5-year journal impact factor (5-Y JIF) is provided by the 2023 Journal Citation Reports (Clarivate) and CiteScores (CS) are provided by SCOPUS. b) Thematic map of the included studies.

exhibited an "unclear" risk of bias in the domain of "patient selection", primarily due to insufficient information regarding the participant selection process, which in turn affected the applicability domain of "patient selection". With respect to the "flow and timing" and "reference standard" domains, 38% and 46% of studies, respectively, were rated as unclear in terms of risk of bias, largely due to incomplete reporting. This lack of clarity also impacted the assessment of applicability, with 43% of studies showing unclear applicability in relation to the reference standard. In contrast, 79% of studies had a low risk of bias for the "index test" and all studies (100%) demonstrated low concerns for the applicability of the index test.
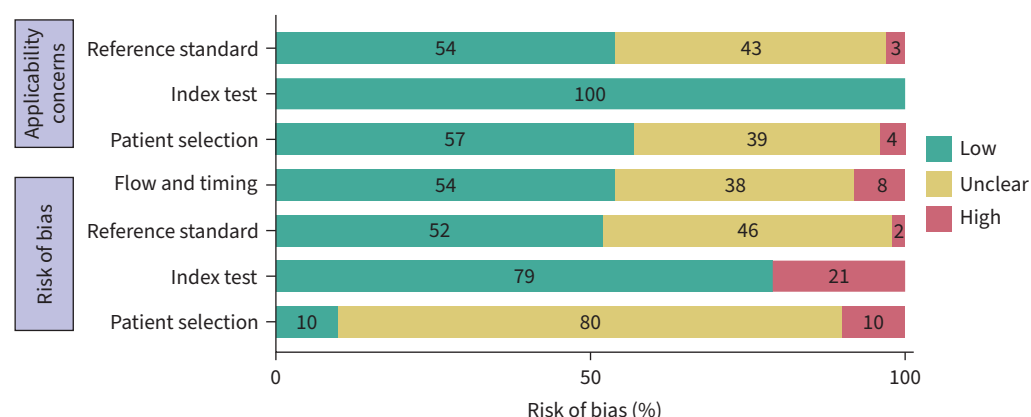
FIGURE 4 Quality assessment of included studies.

Obstructive diseases

A total of 35 studies were noted on obstructive respiratory diseases, with 19 of them concentrating on asthma [36–54] and 16 studies on COPD [55–70] (supplementary table S3).

Asthma
Sound acquisition methods

Studies evaluated stethoscopic lung sounds (n=8) [41, 43–46, 50, 53, 54], non-stethoscopic lung sounds (n=6) [36, 38–40, 49, 52], voluntary or spontaneous cough sounds (n=5) [37, 42, 47, 48, 51], speech (n=1) [36], and vocalised /ɑ:/ sounds (n=1) [37]. The sounds were most frequently recorded through smartphones (n=6) [37, 38, 42, 47, 51, 52], stethoscope-based systems (either digital stethoscopes or combinations of microphones and stethoscope bells) (n=8) [41, 43–46, 50, 53, 54], and external sound sensors and microphones (n=5) [36, 39, 40, 49, 52]. Three studies additionally evaluated the combination of clinical variables (*e.g.* sex, weight, age) and sound data [36, 47, 53].

Data analysis methods

Nearly 75% of studies (n=14) [36–38, 41–47, 51–54] classified patients through ML techniques using a variety of classifiers, particularly with support vector machines (SVM) (n=4) [36, 43, 45, 46] and artificial neural network (ANN)-based methods (n=5) [38, 41, 43, 44, 52]. Over a quarter of studies (n=5) relied on conventional statistical analysis through the use of indexes, ratios, automated cough counts or coughs identified by trained examiners [39, 40, 48–50]. Only one study used non-automated analysis, through the manual identification of coughs by trained examiners [49].

Clinical utility

Most studies focused on the diagnosis of asthma (n=10) [37, 38, 41–44, 48–50, 52]. Other clinical utilities included the detection of asthma exacerbations (n=2) [47, 53], the stratification of disease severity (n=6) [36, 40, 45, 46, 51, 54] and the monitoring of treatment response (n=1) [39].

Among the ten studies focusing on the diagnosis of asthma, 80% (n=8) obtained a balanced set of high sensitivities and specificities, with sensitivity values of 80–96% and specificity values of 83–100% [37, 38, 41–44, 50, 52]. The only study using non-automated analysis through manual cough counts obtained a moderate sensitivity of 69% and a poor specificity of 34% [49].

Two studies evaluated the diagnosis of asthma exacerbation through a combination of different variables [47, 53]. The first study evaluated the combination of cough sounds and clinical features, achieving an AUC of 0.93, a specificity of 95% and a sensitivity of 64% [47]. The second study used smart stethoscope-based data and clinical features, reporting an AUC of up to 0.94 [53].

Of the six studies aiming to stratify asthma severity [36, 40, 45, 46, 51, 54], three of them [45, 46, 51] showed high sensitivities (82–91%) and moderate to high specificities (69–97%) for distinguishing between severe asthma and milder disease. Meanwhile, only one study aimed to monitor patient response to inhaled corticosteroid treatment and observed AUCs between 0.86 and 0.92 for the distinction between patients with well-controlled and not-well-controlled asthma [39].

COPD

Sound acquisition methods

Various physiological sounds such as stethoscopic lung sounds (n=8) [55, 56, 61, 63, 66, 68–70], cough sounds (n=4) [57, 64, 65, 67], non-stethoscopic respiratory sounds (n=3) [59, 60, 62] and speech (n=1) [58] were evaluated in COPD studies. These sounds were most frequently recorded through stethoscope-based systems (n=8) [55, 56, 61, 63, 66, 68–70] followed by smartphones (n=4) [57, 64, 65, 67] and external sound sensors or microphones (n=4) [58–60, 62]. Four studies additionally evaluated clinical variables (*e.g.* age, smoking, presence of acute cough, fever) in combination with sound data [57, 61, 64, 65].

Data analysis methods

Only one study used conventional statistical analysis through qualitative and quantitative assessment of vibration response imaging [60]. The rest (n=15) [55–59, 61–70] employed automated ML techniques such as SVM (n=6) [56, 59, 61, 65, 68, 69], decision trees (n=6) [56, 58, 61, 62, 67, 68] and ANN-based methods (n=6) [55, 56, 66, 68–70], with numerous studies evaluating more than one ML analysis method.

Clinical utility

Most studies (n=10) focused on the diagnosis of COPD [55, 56, 58, 60, 61, 63, 65–67, 70]. Of these, 50% (n=5) demonstrated high diagnostic performances, with a combination of sensitivities of 80–100%, and specificities of 82–100% [55, 56, 61, 66, 67]. Two studies additionally found a moderate set of sensitivities (70–73%) and specificities (70–74.5%) [58, 65]. One study differentiated COPD patients from asthma patients with a sensitivity of 81% and a specificity of 74% [60]. Another study distinguished COPD and congestive heart failure patients from healthy controls with a sensitivity of 80% and a specificity of 82% [67].

For the detection of acute exacerbations of COPD (AECOPD), sensitivities and specificities were 83–100% and 74–91%, respectively [57, 64]. Additionally, three studies that assessed disease control, specifically the prediction of patients at risk of developing AECOPD, found moderate to high sensitivities (74–93%) and high specificities (86–98%) [59, 62, 69].

For COPD severity stratification, one study explored the use of stethoscopic breathing sounds for differentiating patients with moderate to severe COPD from patients with milder disease [68]. The study reported high sensitivities (88%) and specificities (91%), suggesting that this method is effective in differentiating between these levels of severity [68].

Infectious respiratory diseases

A total of 19 articles investigated the use of digital audio biomarkers for the diagnosis of infectious respiratory diseases [71–89] (supplementary table S4). Nearly half of these studies (n=9) focused on the general diagnosis of pneumonia [71, 73–75, 78, 82–84, 87], while two studies aimed at differentiating pneumonia from bronchitis [86, 88] and one study targeted lower respiratory tract infections [85]. Specific aetiologies of infectious respiratory diseases were also investigated, such as tuberculosis (TB) (n=3) [72, 76, 89], pertussis (n=3) [77, 79, 80] and croup (n=1) [81]. Four studies also included the evaluation of clinical factors in combination with sound data [71, 72, 78, 89].

Sound acquisition methods

Most studies were based on voluntary or spontaneous cough sounds (n=14) [71–74, 76–81, 85, 86, 88, 89]. Breathing or tracheal sounds (n=5) were also reported for the detection of infectious respiratory diseases [75, 82–84, 87]. Microphones were used in nearly a third of the studies (n=6) [71–76]. Other studies used various data sources such as pre-recorded sound files from the public domain (n=4) [77, 79, 80, 88], smartphones (n=4) [78, 81, 85, 89] and electronic stethoscopes (n=4) [82–84, 87].

Data analysis methods

ML algorithms including linear regression (n=8) [71, 72, 74, 76, 78–81], SVM (n=8) [76, 80, 81, 83–86, 88] and ANN-based methods (n=8) [73, 75–77, 80, 84, 86, 89] were used for sound analysis in infectious respiratory diseases. Only two studies relied on conventional statistical analysis through manual identification of crackles or wheezes [82, 87].

Clinical utility

Among the eleven studies focusing on the diagnosis of pneumonia, seven of them showed high (80–100%) sensitivities [71, 73, 74, 78, 83, 86, 88] and seven of them demonstrated high specificities [71, 75, 78, 82, 83, 86, 87]. One study achieved a sensitivity of 72% and a specificity of 82% in distinguishing patients with pneumonia and COPD from those with a non-pneumonic AECOPD based on stethoscopic tracheal

sounds [75]. One study distinguished pneumonia patients from bronchitis patients with a sensitivity of 93% and a specificity of 89% [86]. Three studies focused on TB diagnosis and reported high sensitivities (82–93%) and specificities (81–95%) [72, 76, 89]. Three studies evaluated pertussis using cough recordings from publicly available sources and obtained excellent sensitivities of 93–100% and specificities of 92–100% [77, 79, 80]. For croup diagnosis, only one study was reported to use cough sounds, with a sensitivity of 92% and a specificity of 85% [81].

Various respiratory diseases

A total of 27 records were included in the category of various respiratory diseases and addressed obstructive or infectious conditions but were not limited exclusively to one of these categories (supplementary table S5) [90–116].

Sound acquisition methods

The included studies most frequently evaluated stethoscopic lung sounds (n=19) [93, 94, 99, 101–116] and voluntary or spontaneous cough sounds (n=5) [90–92, 97, 98], while fewer studies (n=3) [95, 96, 100] focused on non-stethoscopic lung sounds. The sounds were most frequently recorded through stethoscope-based systems (n=19) [93, 94, 99, 101–116], non-stethoscope-based microphones or sensors (n=6) [90, 92, 95, 96, 98, 100] and smartphones (n=2) [91, 97]. Approximately 60% of studies (n=16) used sound files obtained from a publicly available dataset, such as the ICBHI 2017 (n=16) [93, 94, 99, 102–104, 106, 108–116], the Jordan University of Science and Technology dataset (n=3) [103, 104, 106] and others [109, 111, 115] (supplementary figure S1).

Data analysis methods

For studies relying on conventional statistical models (n=3), variables such as end-point crackling, beginning of crackling or diagnosis by experienced listeners were used for analysis [92, 95, 100]. Few studies (n=2) used non-automated data analysis methods [92, 95]. The vast majority of studies (n=24, 88.9%) used ML classifiers, among which decision trees (n=7) [93, 98, 102, 108–110, 115] and ANN-based methods, such as general ANN methods (n=1) [109], recurrent neural networks (n=5) [91, 104, 111, 114, 116], convolutional neural networks (n=11) [94, 99, 104–107, 111, 112, 114–116], feed-forward neural networks (n=1) [98], specialised neural networks (n=1) [113] and time-delay neural networks (n=1) [97], were the most popular. Many studies investigated more than one ML analysis method.

Clinical utility

Approximately 65% of studies (n=18) compared patients with various respiratory diseases (including asthma, COPD, infections, bronchiectasis, *etc*.) and healthy controls with a combination of high sensitivities (87–99%) and specificities (84–99%) [90, 91, 93–95, 98, 99, 101–104, 106–111, 113]. The respiratory diseases groups evaluated in studies included in this category were extremely heterogeneous, including conditions such as asthma (n=19) [91–93, 95–99, 102–107, 109–111, 113, 114], COPD (n=22) [93–96, 98–106, 108–116], pneumonia (n=23) [92–94, 97–116], bronchiolitis (n=15) [92, 94, 97, 99, 102, 104, 107–109, 111–116], bronchiectasis (n=17) [93, 94, 98–100, 102–105, 109–116], upper respiratory tract infection (n=16) [91, 94, 97, 99, 102, 104, 106, 108–116], lower respiratory tract infections (n=11) [91, 97, 99, 102, 104, 106, 109–111, 113, 114], obstructive or restrictive lung diseases not otherwise specified (n=3) [90, 95, 98], fibrosing alveolitis (n=2) [96, 100], interstitial lung disease (n=2) [98, 105], pulmonary fibrosis (n=3) [101, 106, 111], bronchitis (n=5) [93, 98, 103, 104, 106], heart failure (n=6) [93, 100, 101, 103, 104, 106], pleural effusion (n=1) [106], croup (n=1) [97] and tuberculosis (n=1) [98].

Study data synthesis and critical appraisal

In this review, we identified promising diagnostic accuracies of audio-based digital biomarkers for the diagnosis and management of various respiratory diseases (figure 5). The best diagnostic accuracies were observed in asthma diagnosis, with 80% of studies (n=8 out of 10) reporting a favourable combination of both high sensitivities and specificities (⩾80%). Similarly, for distinguishing COPD from healthy controls, 62.5% of studies (n=5 out of 8) reported a combination of high sensitivities and specificities (⩾80%). In the case of pneumonia, a combination of high sensitivities and specificities (⩾80%) was observed in 36.4% of studies (n=4 out of 11), while 63.6% (n=7 out of 11) demonstrated high sensitivities and 63.6% (n=7 out of 11) demonstrated high specificities. While most studies focused on disease diagnosis, a handful of studies explored the use of acoustic biomarkers to evaluate the presence of disease exacerbations (asthma exacerbation n=2, AECOPD n=2), disease severity (asthma n=6, COPD n=1), disease control (COPD n=3) and treatment response (asthma n=1). The constrained number of studies exploring these additional objectives limits the conclusions that can be drawn regarding their accuracies.
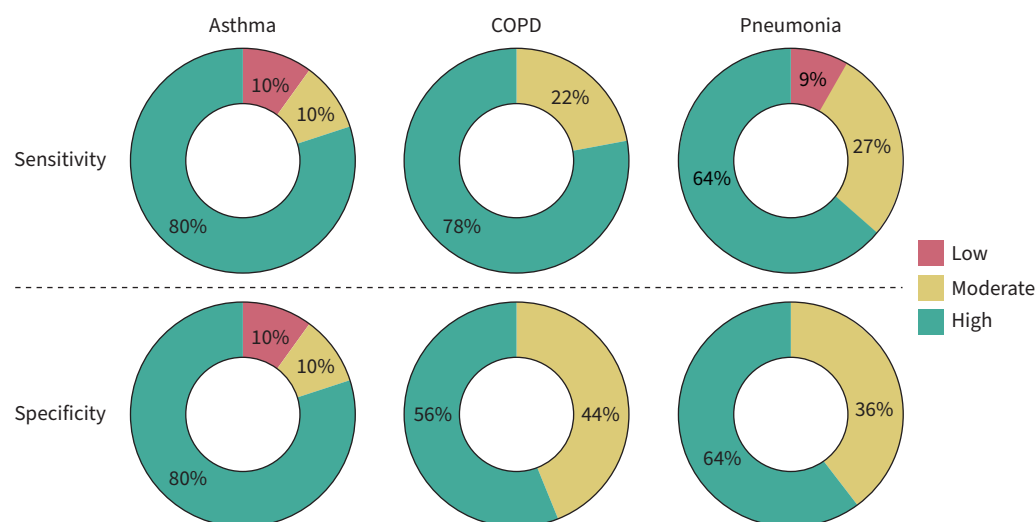
FIGURE 5 Diagnostic accuracies of audio-based digital biomarkers for the diagnosis and management of asthma, COPD and pneumonia. Green indicates high accuracies (80–100%), yellow indicates moderate accuracies (60–79%) and red indicates low accuracies (0–59%).

Study data further revealed that breathing sounds were the most commonly analysed audio data, particularly in studies involving asthma and COPD (supplementary figure S2). Cough sounds were notably more prevalent in studies focusing on infectious diseases because cough is often a primary symptom of these conditions. In contrast, obstructive diseases like asthma or COPD tend to present more frequently with dyspnoea or wheezing as their primary symptoms. Furthermore, cough sounds have long played a pivotal role in distinguishing between various infectious respiratory conditions, such as the characteristic whooping cough of pertussis or the seal bark cough in croup. This likely accounts for the increased focus on cough sounds in studies evaluating infectious diseases and is reflected by the presence of cough-related keywords as motor themes in figure 3b. The inclusion of speech sounds and vocalised sounds remained minimal across the included studies. Besides stethoscope-based devices, mobile phones and external microphones have remained the principal data acquisition devices in the last 5 years (supplementary figure S3).

For analytical methods, ANN-based methods were the most frequently used, especially in the "Various diseases" category. Studies using ANN were often more exploratory in nature, aiming to assess and compare multiple analytical methods (supplementary figure S4). SVM, linear regression and decision tree methods were also frequently used across the different conditions.

Only a small subset of studies (n=8 out of 81) investigated data acquisition in home or workplace settings, reflecting the early, pre-clinical stage of research in this field and potential research logistical challenges [48, 49, 53, 58, 59, 62, 69, 95]. Due to this small number of studies, it is difficult to draw meaningful comparisons or conclusions regarding potential differences in accuracy or validity between settings.

Discussion

Significance within current literature

To the best of our knowledge, this study represents the first comprehensive review of this scale, directed toward the characterisation of the accuracy of sound analysis in the diagnosis and management of prevalent airway diseases. This review is also the first to evaluate an umbrella of audio biomarkers, including cough, breathing sounds, speech, and vocalised sounds.

Previously published reviews primarily focused on the detection of abnormal lung sounds, as opposed to the diagnosis of diseases as included here. For instance, REICHERT et al. [13] and HEGDE et al. [15] conducted reviews exploring the use of cough sounds as screening and diagnostic tools, but did not specifically address diagnostic accuracies. GURUNG et al. [14] conducted a systematic review and meta-analysis on computerised lung sound analysis, which focused on the identification of abnormal lung sounds such as wheezes or crackles as opposed to specific clinical diagnoses of respiratory disease. Similarly, PALANIAPPAN et al. [18, 118] and PRAMONO et al. [119] published systematic reviews on lung sound analysis, which did not focus on the diagnosis of respiratory diseases.

Beyond differentiating between abnormal and normal lung sounds, this paper further contributed to reviewing the diagnostic utility of audio-based measures. Harnessing physiological data from multiple modalities and sources such as respiratory rates, body temperature, sleeping patterns and physiological sounds has been a challenge in the past owing to hardware and software limitations. With the emergence of multi-modal wearable devices and deep-learning algorithms, combining audio-based and other respiratory measures for clinical diagnosis is a promising development in advancing precision medicine for respiratory diseases [120].

Challenges and considerations for the use of audio biomarkers and ML in clinical settings

While physiological sound analysis holds potential for diagnosing asthma, COPD and infectious respiratory disorders, notable challenges remain in translating these techniques into clinical routines. Along the pipeline, the first step of physiological sound analysis is acoustic data collection, which can be achieved with digital stethoscopes, microphones, smartphones, *etc*. [6, 13, 121]. The characteristics of the sound recording systems, including their size, sensor quantity and installation procedures, significantly impact their clinical utility. Complex sound-recording systems (*e.g.* those requiring the installation of multiple sensors in hard-to-reach locations) may require installation by a trained professional; meanwhile, sound-recording systems such as smartphones are highly dependent on the patient's own technological proficiency. Notably, the sound acquisition step is particularly vulnerable to the quality of available equipment and to the environmental conditions in which sound capture takes place [121–123]. In busy clinical settings, a high level of ambient noise will lead to subpar signal quality for accurate downstream data analysis [121–123].

Noise reduction and audio filtering techniques have been described to minimise acoustic artefacts [13, 121, 123, 124]. These methods suppress unwanted noise within an audio sample while preserving the desired audio content. Audio filters suppress or enhance frequencies within a certain range [124]. Other methods involve adaptive noise cancellation, spectral restoration techniques and machine learning-based algorithms [13, 124]. Ultimately, the optimal noise reduction technique depends on a variety of factors such as the characteristics of the desired audio content (*e.g.* speech *versus* stethoscopic lung sounds), unwanted noise (*e.g.* background speech *versus* hums or heart sounds), desired quality of resulting audio sample and resources available [13, 124].

After noise reduction and filtration, the audio signal is typically divided into short overlapping segments through frame segmentation and then undergoes feature extraction [122, 123, 125]. Features are usually combined into a feature vector and then subjected to machine-aided classification analysis. This review revealed that the most-used classifiers include random forest, linear regression, SVM, k-nearest neighbour and ANN in respiratory disease detection. Classifier selection is dependent on data and feature types. For example, random forest and linear regression are both used to predict outcomes in one of two categories (*e.g.* yes/no). Random forest models tend to offer high prediction accuracies and are easily understandable but have been shown to work best for datasets with fewer features [36]. For more complex data, ANNs offer an advantage in their ability to detect nonlinear relationships [126]. Within neural networks, relational neural networks are promising for predictive applications in speech and language, where data is often structured as a sequential stream [127]. When fewer input data are involved, feature-based ML classifiers such as random forest and SVM usually achieve better performance and explainable classification [128]. However, with data that are more complex, deep-learning methods such as neural networks can better capture complicated relationships between features of the input audio and the output labels, leading to better overall performance. It is worth noting that deep-learning methods often require substantially greater amounts of data (*e.g.* millions of data points) and require significantly more computational power to run, although they often show higher accuracies in the diagnosis of respiratory illnesses [129, 130]. Lower computational constraints on simpler ML classifiers (which often run in minutes, while large deep-learning models can take hours or days) may make them better candidates for home-monitoring contexts, given that remote and digital healthcare practice has become increasingly important in recent years [131]. Notably, computing performance was almost never reported in the included studies but constitutes an important consideration when integrating this technology into resource-limited mobile devices.

To date, most of these ML-driven methods have primarily been used in research settings and have yet to transition to clinical practice. AI-based analyses are subject to the same challenges in lung sound analysis as they are in many other healthcare domains [131]. Notably, the quality of these algorithms is strongly dependent on the quality of the data used to train them [131]. Samples in the included studies were often small, and poorly characterised. Consequently, the potential for generalising these findings to broader populations remains uncertain. Although speech sound may help improve diagnostic accuracies over the use of physiological sound (*e.g.* coughs, wheezes) in isolation, it also introduces more significant privacy

issues with respect to patient identification. Data protection is especially important in a clinical context owing to potential impacts on employment, insurance and more [132]. Introducing measures such as encryption, and using reporting guidelines and risk-of-bias tools specific to AI or deep-learning clinical trials, such as Consolidated Standards of Reporting Trials-AI (CONSORT-AI), Prediction Model Risk Of Bias Assessment Tool-AI (PROBAST-AI) and Transparent Reporting of a Multivariable Prediction Model of Individual Prognosis Or Diagnosis-AI (TRIPOD-AI), enables researchers to enact the Findability, Accessibility, Interoperability and Reusability (FAIR) principles of AI analysis and facilitate its transition to clinical implementation [133].

Study limitations

First, it is crucial to recognise the important heterogeneity among included studies. This heterogeneity manifests in various aspects, including the diversity of sample populations, ranging from paediatric to adult patients and encompassing variations in disease severity; the type of sound recorded, *e.g.* cough, stethoscopic lung sounds and speech; the wide array of devices used, ranging from complex stethoscope-based systems to smartphones; the type of features extracted, *e.g.* spectral, cepstral and temporal; and the study goals, *e.g.* diagnosis, monitoring and staging. The results of these studies have been grouped according to clinical application (*e.g.* diagnosis, monitoring, staging) and presented together for the purpose of this review. However, highly variable methodologies in sound data acquisition and data analysis methodologies are noted in the study data. As such, a meta-analysis of the literature data was not possible. Second, many studies did not use gold standard methods to establish diagnoses (*e.g.* symptom-based diagnosis of pneumonia), which can limit comparisons across studies and the extent to which definitive conclusions can be drawn regarding their accuracies. Lastly, many of the included studies presented unclear or high concerns for bias and applicability, which highlights a compelling need for higher quality studies that adhere to rigorous protocols for patient selection and reporting.

Conclusion

The present review on physiological sound and speech analysis sheds light on the potential of this noninvasive, contact-free approach for the diagnosis and management of common respiratory diseases. The current body of evidence consistently demonstrated moderate (60–79%) to high (80–100%) accuracies for the diagnosis of COPD, asthma, and infectious respiratory diseases across a wide variety of sound analysis techniques as well as a wide array of devices and feature extraction methods in pre-clinical settings. Our review underscores the need for more rigorous study designs, better characterisation of study populations, and the use of gold standard diagnostic criteria to facilitate meaningful comparisons across studies. Clinically diverse populations can be further included in the research protocol to enhance clinical validity. Data security, algorithmic bias and user privacy remain significant challenges in the development of AI-driven diagnostic tools using sound-based data. A rigorous implementation of comprehensive standards and the development of robust regulation frameworks (*e.g.* TRIPOD-AI) is necessary for addressing algorithmic bias and promoting trust and transparency of AI in respiratory medicine. With continued advances in the field of bioacoustics and machine intelligence, physiological sound technology may provide a complementary tool to existing clinical tools in the diagnosis and management of respiratory diseases.

Questions for future research

Recent years have seen significant advances in audio-based digital biomarkers for respiratory diagnostics, largely driven by improvements in portable data collection devices, and AI-based data analysis methods. However, numerous questions and challenges must be addressed for safe, efficient and appropriate clinical implementation.

1. How will audio-based digital biomarkers contribute to and enhance the multimodal diagnosis of respiratory diseases when combined with other variables such as clinical data and other biomarkers?
2. Will the same audio-based digital biomarkers validated for diagnostic use be equally accurate for other applications, such as remote monitoring of disease progression, remission or exacerbation?
3. In the context of audio data acquisition, particularly in private settings or for prolonged periods of time, what gold standard measures should be taken to ensure data safety and confidentiality?

References

1   Forum of International Respiratory Societies (FIRS). The Global Impact of Respiratory Disease - 3rd Edition. Sheffield, European Respiratory Society, 2021.

2   Price D, Freeman D, Cleland J, *et al.* Earlier diagnosis and earlier treatment of COPD in primary care. *Prim Care Respir J* 2011; 20: 15–22.

3   Bohadana A, Izbicki G, Kraman SS. Fundamentals of lung auscultation. *N Engl J Med* 2014; 370: 744–751.

4   Sarkar M, Madabhavi I, Niranjan N, *et al.* Auscultation of the respiratory system. *Ann Thorac Med* 2015; 10: 158–168.

5   Hafke-Dys H, Bręborowicz A, Kleka P, *et al.* The accuracy of lung auscultation in the practice of physicians and medical students. *PLoS One* 2019; 14: e0220606.

6   Kim Y, Hyon Y, Lee S, *et al.* The coming era of a new auscultation system for analyzing respiratory sounds. *BMC Pulm Med* 2022; 22: 119.

7   Arts L, Lim EHT, van de Ven PM, *et al.* The diagnostic accuracy of lung auscultation in adult patients with acute pulmonary pathologies: a meta-analysis. *Sci Rep* 2020; 10: 7347.

8   Mangione S, Nieman LZ. Pulmonary auscultatory skills during training in internal medicine and family practice. *Am J Respir Crit Care Med* 1999; 159: 1119–1124.

9   Meghji J, Mortimer K, Agusti A, *et al.* Improving lung health in low-income and middle-income countries: from challenges to solutions. *Lancet* 2021; 397: 928–940.

10   Deesha P, Ajiri A, Nyasha C, *et al.* Attitudes to participation in a lung cancer screening trial: a qualitative study. *Thorax* 2012; 67: 418–425.

11   Lukas H, Xu C, Yu Y, *et al.* Emerging telemedicine tools for remote COVID-19 diagnosis, monitoring, and management. *ACS Nano* 2020; 14: 16180–16193.

12   Gupta P, Wen H, Di Francesco L, *et al.* Detection of pathological mechano-acoustic signatures using precision accelerometer contact microphones in patients with pulmonary disorders. *Sci Rep* 2021; 11: 13427.

13   Reichert S, Gass R, Brandt C, *et al.* Analysis of respiratory sounds: state of the art. *Clin Med Circ Respirat Pulm Med* 2008; 2: 45–58.

14   Gurung A, Scrafford CG, Tielsch JM, *et al.* Computerized lung sound analysis as diagnostic aid for the detection of abnormal lung sounds: a systematic review and meta-analysis. *Respir Med* 2011; 105: 1396–1403.

15   Hegde S, Sreeram S, Alter IL, *et al.* Cough sounds in screening and diagnostics: a scoping review. *Laryngoscope* 2024; 134: 1023-1031.

16   Lei Z, Martignetti L, Ridgway C, *et al.* Wearable neck surface accelerometers for occupational vocal health monitoring: instrument and analysis validation study. *JMIR Form Res* 2022; 6: e39789.

17   Lei Z, Kennedy E, Fasanella L, *et al.* Discrimination between modal, breathy and pressed voice for single vowels using neck-surface vibration signals. *Appl Sci* 2019; 9:1505.

18   Palaniappan R, Sundaraj K, Ahamed NU. Machine learning in lung sound analysis: a systematic review. *Biocybernet Biomed Eng* 2013; 33:129–135.

19   Chamberlain D, Kodgule R, Ganelin D, *et al.* Application of semi-supervised deep learning to lung sound analysis. *In*: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 16–20 August 2016, 2016.

20   Barry SJ, Dane AD, Morice AH, *et al.* The automatic recognition and counting of cough. *Cough* 2006; 2: 8.

21   Ng AK, Koh TS, Baey E, *et al.* Could formant frequencies of snore signals be an alternative means for the diagnosis of obstructive sleep apnea? *Sleep Med* 2008; 9: 894–898.

22   Groh R, Lei Z, Martignetti L, *et al.* Efficient and explainable deep neural networks for airway symptom detection in support of wearable health technology. *Adv Intell Syst* 2022; 4: 2100284.

23   Moher D, Liberati A, Tetzlaff J, *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 2010; 8: 336–341.

24   McGowan J, Sampson M, Salzwedel DM, *et al.* PRESS peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol* 2016; 75: 40–46.

25   Geersing G-J, Bouwmeester W, Zuithoff P, *et al.* Search filters for finding prognostic and diagnostic prediction studies in MEDLINE to enhance systematic reviews. *PLoS One* 2012; 7: e32844.

26   Haynes RB, Wilczynski N, McKibbon KA, *et al.* Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *J Am Med Inform Assoc* 1994; 1: 447–458.

27   Eriksen MB, Frandsen TF. The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review. *J Med Libr Assoc* 2018; 106: 420–431.

28   Page MJ, Moher D, Bossuyt PM, *et al.* PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021; 372: n160.

29   Whiting PF, Rutjes AW, Westwood ME, *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155: 529–536.

30   Aria M, Cuccurullo C. bibliometrix: an R-tool for comprehensive science mapping analysis. *J Informetr* 2017; 11: 959–975.

31   Pereira V, Basilio M, Tarjano C. pyBibX -- A Python Library for Bibliometric and Scientometric Analysis Powered with Artificial Intelligence Tools. arXiv 2023; preprint. [https://doi.org/10.48550/arXiv.2304.14516]

32   Clarivate. Journal Citation Reports: How to Cite Editions. 2022. https://support.clarivate.com/Scientificand AcademicResearch/s/article/Journal-Citation-Reports-How-to-Cite-Editions?language=en_US Date last accessed: 10 January 2025.

33   Mühl DD, de Oliveira L. A bibliometric and thematic approach to agriculture 4.0. *Heliyon* 2022; 8: e09369.

34   Cobo MJ, Martínez MA, Gutiérrez-Salcedo M, *et al.* 25 years at Knowledge-Based Systems: a bibliometric analysis. *Knowl Based Syst* 2015; 80: 3–13.

35   Ouzzani M, Hammady H, Fedorowicz Z, *et al.* Rayyan – a web and mobile app for systematic reviews. *Syst Rev* 2016; 5: 210.

36   Alam MZ, Simonetti A, Brillantino R, *et al.* Predicting pulmonary function from the analysis of voice: a machine learning approach. *Front Digit Health* 2022; 4: 750226.

37   B TB, Hee HI, Teoh OH, *et al.* Asthmatic versus healthy child classification based on cough and vocalised /ɑ:/ sounds. *J Acoust Soc Am* 2020; 148: El253.

38   Gelman A, Furman EG, Kalinina NM, *et al.* Computer-aided detection of respiratory sounds in bronchial asthma patients based on machine learning method. *Sovrem Tekhnologii Med* 2022; 14: 45–51.

39   Habukawa C, Murakami K, Endoh M, *et al.* Treatment evaluation using lung sound analysis in asthmatic children. *Respirology* 2017; 22: 1564–1569.

40   Habukawa C, Murakami K, Horii N, *et al.* A new modality using breath sound analysis to evaluate the control level of asthma. *Allergol Int* 2013; 62: 29–35.

41   Hafke-Dys H, Kuźnar-Kamińska B, Grzywalski T, *et al.* Artificial intelligence approach to the monitoring of respiratory sounds in asthmatic patients. *Front Physiol* 2021; 12: 745635.

42   Hee HI, Balamurali BT, Karunakaran A, *et al.* Development of machine learning for asthmatic and healthy voluntary cough sounds: a proof of concept study. *Appl Sci* 2019; 9: 2833.

43   Islam MA, Bandyopadhyaya I, Bhattacharyya P, *et al.* Multichannel lung sound analysis for asthma detection. *Comput Methods Programs Biomed* 2018; 159: 111–123.

44   Lin BS, Wu HD, Chen SJ. Automatic wheezing detection based on signal processing of spectrogram and back-propagation neural network. *J Healthc Eng* 2015; 6: 649–672.

45   Nabi FG, Sundaraj K, Lam CK, *et al.* Characterization and classification of asthmatic wheeze sounds according to severity level using spectral integrated features. *Comput Biol Med* 2019; 104: 52–61.

46   Nabi FG, Sundaraj K, Lam CK. Identification of asthma severity levels through wheeze sound characterization and classification using integrated power features. *Biomed Signal Process Control* 2019; 52: 302–311.

47   Porter P, Brisbane J, Abeyratne U, *et al.* A smartphone-based algorithm comprising cough analysis and patient-reported symptoms identifies acute exacerbations of asthma: a prospective, double blind, diagnostic accuracy study. *J Asthma* 2023; 60: 368–376.

48   Rhee H, Belyea MJ, Sterling M, *et al.* Evaluating the validity of an automated device for asthma monitoring for adolescents: correlational design. *J Med Internet Res* 2015; 17: e234.

49   Rietveld S, Rijssenbeek-Nouwens LH. Diagnostics of spontaneous cough in childhood asthma: results of continuous tracheal sound recording in the homes of children. *Chest* 1998; 113: 50–54.

50   Shimoda T, Nagasaka Y, Obase Y, *et al.* Prediction of airway inflammation in patients with asymptomatic asthma by using lung sound analysis. *J Allergy Clin Immunol Pract* 2014; 2: 727–732.

51   Swarnkar V, Abeyratne U, Tan J, *et al.* Stratifying asthma severity in children using cough sound analytic technology. *J Asthma* 2021; 58: 160–169.

52   Aptekarev T, Sokolovsky V, Furman E, *et al.* Application of deep learning for bronchial asthma diagnostics using respiratory sound recordings. *PeerJ Comput Sci* 2023; 9: e1173.

53   Emeryk A, Derom E, Janeczek K, *et al.* Home monitoring of asthma exacerbations in children and adults with use of an AI-aided stethoscope. *Ann Fam Med* 2023; 21: 517–525.

54   Ghulam Nabi F, Sundaraj K, Iqbal MS, *et al.* A telemedicine software application for asthma severity levels identification using wheeze sounds classification. *Biocybernet Biomed Eng* 2022; 42: 1236–1247.

55   Altan G, Kutlu Y, Pekmezci AÖ, *et al.* Deep learning with 3D-second order difference plot on respiratory sounds. *Biomed Signal Process Control* 2018; 45: 58–69.

56   Altan G, Kutlu Y, Allahverdi N. Deep learning on computerized analysis of chronic obstructive pulmonary disease. *IEEE J Biomed Health Inform* 2020; 24: 1344–1350.

57   Claxton S, Porter P, Brisbane J, *et al.* Identifying acute exacerbations of chronic obstructive pulmonary disease using patient-reported symptoms and cough feature analysis. *NPJ Digit Med* 2021; 4: 107.

58   Farrús M, Codina-Filbà J, Reixach E, *et al.* Speech-based support system to supervise chronic obstructive pulmonary disease patient status. *Appl Sci* 2021; 11: 7999.

59   Fernandez-Granero MA, Sanchez-Morillo D, Leon-Jimenez A. Computerised analysis of telemonitored respiratory sounds for predicting acute exacerbations of COPD. *Sensors (Basel)* 2015; 15: 26978–26996.

60   Guntupalli KK, Reddy RM, Loutfi RH, *et al.* Evaluation of obstructive lung disease with vibration response imaging. *J Asthma* 2008; 45: 923–930.

61   Haider NS, Singh BK, Periyasamy R, *et al.* Respiratory sound based classification of chronic obstructive pulmonary disease: a risk stratification approach in machine learning paradigm. *J Med Syst* 2019; 43: 255.

62   Fernandez-Granero MA, Sanchez-Morillo D, Leon-Jimenez A. An artificial intelligence approach to early predict symptom-based exacerbations of COPD. *Biotechnol Biotechnol Equip* 2018; 32: 778–784.

63   Naqvi SZ, Choudhry MA. An automated system for classification of chronic obstructive pulmonary disease and pneumonia patients using lung sound analysis. *Sensors (Basel)* 2020; 20: 6512.

64   Porter P, Claxton S, Brisbane J, *et al.* Diagnosing chronic obstructive airway disease on a smartphone using patient-reported symptoms and cough analysis: diagnostic accuracy study. *JMIR Form Res* 2020; 4: e24587.

65   Sharan RV, Abeyratne UR, Swarnkar VR, *et al.* Predicting spirometry readings using cough sound features and regression. *Physiol Meas* 2018; 39: 095001.

66   Srivastava A, Jain S, Miranda R, *et al.* Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease. *PeerJ Comput Sci* 2021; 7: e369.

67   Windmon A, Minakshi M, Bharti P, *et al.* TussisWatch: a smart-phone system to identify cough episodes as early symptoms of chronic obstructive pulmonary disease and congestive heart failure. *IEEE J Biomed Health Inform* 2019; 23: 1566–1573.

68   Yu H, Zhao J, Liu D, *et al.* Multi-channel lung sounds intelligent diagnosis of chronic obstructive pulmonary disease. *BMC Pulm Med* 2021; 21: 321.

69   Yin H, Wang K, Yang R, *et al.* A machine learning model for predicting acute exacerbation of in-home chronic obstructive pulmonary disease patients. *Comput Methods Programs Biomed* 2024; 246: 108005.

70   Le Trung K, Nguyen Anh P, Han T-T. A novel method in COPD diagnosing using respiratory signal generation based on CycleGAN and machine learning. *Comput Methods Biomech Biomed Engin* 2024; in press [https://doi.org/10.1080/10255842.2024.2329938].

71   Abeyratne UR, Swarnkar V, Setyati A, *et al.* Cough sound analysis can rapidly diagnose childhood pneumonia. *Ann Biomed Eng* 2013; 41: 2448–2462.

72   Botha GHR, Theron G, Warren RM, *et al.* Detection of tuberculosis by automatic cough sound analysis. *Physiol Meas* 2018; 39: 045005.

73   Chung Y, Jin J, Jo HI, *et al.* Diagnosis of pneumonia by cough sounds analyzed with statistical features and AI. *Sensors (Basel)* 2021; 21: 7036.

74   Kosasih K, Abeyratne UR, Swarnkar V, *et al.* Wavelet augmented cough analysis for rapid childhood pneumonia diagnosis. *IEEE Trans Biomed Eng* 2015; 62: 1185–1194.

75   Morillo DS, León Jiménez A, Moreno SA. Computer-aided diagnosis of pneumonia in patients with chronic obstructive pulmonary disease. *J Am Med Inform Assoc* 2013; 20: e111–e117.

76   Pahar M, Klopper M, Reeve B, *et al.* Automatic cough classification for tuberculosis screening in a real-world environment. *Physiol Meas* 2021; 42: 105014.

77   Parker D, Picone J, Harati A, *et al.* Detecting paroxysmal coughing from pertussis cases using voice recognition technology. *PLoS One* 2014; 8: e82971.

78   Porter P, Brisbane J, Abeyratne U, *et al.* Diagnosing community-acquired pneumonia *via* a smartphone-based algorithm: a prospective cohort study in primary and acute-care consultations. *Br J Gen Pract* 2021; 71: e258–e265.

79   Pramono RXA, Imtiaz SA, Rodriguez-Villegas E. A cough-based algorithm for automatic diagnosis of pertussis. *PLoS One* 2016; 11: e0162128.

80   Sharan RV, Berkovsky S, Navarro DF, *et al.* Detecting pertussis in the pediatric population using respiratory sound events and CNN. *Biomed Signal Processing Control* 2021; 68: 102722.

81   Sharan RV, Abeyratne UR, Swarnkar VR, *et al.* Automatic croup diagnosis using cough sound recognition. *IEEE Trans Biomed Eng* 2019; 66: 485–495.

82   Scrafford CG, Basnet SC, Ansari I, *et al.* Evaluation of digital auscultation to diagnose pneumonia in children 2 to 35 months of age in a clinical setting in Kathmandu, Nepal: a prospective case–control study. *J Pediatr Infect Dis* 2016; 11: 028–036.

83   Haider NS, Behera AK. Computerized respiratory sound based diagnosis of pneumonia. *Med Biol Eng Comput* 2024; 62: 95–106.

84   Huang D, Wang L, Wang W. A multi-center clinical trial for wireless stethoscope-based diagnosis and prognosis of children community-acquired pneumonia. *IEEE Trans Biomed Eng* 2023; 70: 2215–2226.

85   Yuling L, Siyu Y, Zaiting Y, *et al.* Identification of elderly patients with lower respiratory tract infection by artificial intelligence analysis of cough pattern sounds. *Discov Med* 2023; 35: 1160–1166.

86   Liao S, Song C, Wang X, *et al.* A classification framework for identifying bronchitis and pneumonia in children based on a small-scale cough sounds dataset. *PLoS One* 2022; 17: e0275479.

87   Park DE, Watson NL, Focht C, *et al.* Digitally recorded and remotely classified lung auscultation compared with conventional stethoscope classifications among children aged 1–59 months enrolled in the Pneumonia Etiology Research for Child Health (PERCH) case–control study. *BMJ Open Respir Res* 2022; 9: e001144.

88   Sharan RV, Qian K, Yamamoto Y. Automated cough sound analysis for detecting childhood pneumonia. *IEEE J Biomed Health Inform* 2024; 28: 193–203.

89   Yellapu GD, Rudraraju G, Sripada NR, *et al.* Development and clinical validation of Swaasa AI platform for screening and prioritization of pulmonary TB. *Sci Rep* 2023; 13: 4740.

90   Abaza AA, Day JB, Reynolds JS, *et al.* Classification of voluntary cough sound and airflow patterns for detecting abnormal pulmonary function. *Cough* 2009; 5: 8.

91   Balamurali BT, Hee HI, Kapoor S, *et al.* Deep neural network-based respiratory pathology classification using cough sounds. *Sensors (Basel)* 2021; 21: 5555.

92   Bisballe-Müller N, Chang AB, Plumb EJ, *et al.* Can acute cough characteristics from sound recordings differentiate common respiratory illnesses in children? A comparative prospective study. *Chest* 2021; 159: 259–269.

93   Fraiwan L, Hassanin O, Fraiwan M, *et al.* Automatic identification of respiratory diseases from stethoscopic lung sound signals using ensemble classifiers. *Biocybernet Biomed Eng* 2021; 41: 1–14.

94   García-Ordás MT, Benítez-Andrades JA, García-Rodríguez I, *et al.* Detecting respiratory pathologies using convolutional neural networks and variational autoencoders for unbalancing data. *Sensors* 2020; 20: 1214.

95   Gavriely N, Nissan M, Cugell DW, *et al.* Respiratory health screening using pulmonary function tests and lung sound analysis. *Eur Respir J* 1994; 7: 35–42.

96   Malmberg LP, Kallio K, Haltsonen S, *et al.* Classification of lung sounds in patients with asthma, emphysema, fibrosing alveolitis and healthy lungs by using self-organizing maps. *Clin Physiol* 1996; 16: 115–129.

97   Porter P, Abeyratne U, Swarnkar V, *et al.* A prospective multicentre study testing the diagnostic accuracy of an automated cough sound centred analytic system for the identification of common respiratory disorders in children. *Respir Res* 2019; 20: 81.

98   Rudraraju G, Palreddy S, Mamidgi B, *et al.* Cough sound analysis and objective correlation with spirometry and clinical diagnosis. *Inform Med Unlocked* 2020; 19: 100319.

99   Shuvo SB, Ali SN, Swapnil SI, *et al.* A lightweight CNN model for detecting respiratory diseases from lung auscultation sounds using EMD-CWT-based hybrid scalogram. *IEEE J Biomed Health Inform* 2021; 25: 2595–2603.

100   Sovijärvi AR, Piirilä P, Luukkonen R. Separation of pulmonary disorders with two-dimensional discriminant analysis of crackles. *Clin Physiol* 1996; 16: 171–181.

101   Stasiakiewicz P, Dobrowolski AP, Targowski T, *et al.* Automatic classification of normal and sick patients with crackles using wavelet packet decomposition and support vector machine. *Biomed Signal Process Control* 2021; 67: 102521.

102   Abera Tessema B, Nemomssa HD, Lamesgin Simegn G. Acquisition and classification of lung sounds for improving the efficacy of auscultation diagnosis of pulmonary diseases. *Med Devices (Auckl)* 2022; 15: 89–102.

103   Abdul Sattar Shaikh A, Bhargavi MS, Kumar CP. Weighted aggregation through probability based ranking: an optimized federated learning architecture to classify respiratory diseases. *Comput Methods Programs Biomed* 2023; 242: 107821.

104   Alqudah AM, Qazan S, Obeidat YM. Deep learning models for detecting respiratory pathologies from raw lung auscultation sounds. *Soft Comput* 2022; 26: 13405–13429.

105   Choi Y, Lee H. Interpretation of lung disease classification with light attention connected module. *Biomed Signal Process Control* 2023; 84: 104695.

106   Hassan U, Singhal A, Chaudhary P. Lung disease detection using EasyNet. *Biomed Signal Process Control* 2024; 91: 105944.

107   Heitmann J, Glangetas A, Doenz J, *et al.* DeepBreath – automated detection of respiratory pathology from lung auscultation in 572 pediatric outpatients across 5 countries. *NPJ Digit Med* 2023; 6: 104.

108 Kababulut FY, Gürkan Kuntalp D, Düzyel O, *et al.* A new Shapley-based feature selection method in a clinical decision support system for the identification of lung diseases. *Diagnostics (Basel)* 2023; 13: 3558.

109 Karaarslan O, Belcastro KD, Ergen O. Respiratory sound-base disease classification and characterization with deep/machine learning techniques. *Biomed Signal Process Control* 2024; 87: 105570.

110 Mahmood AF, Alkababji AM, Daood A. Resilient embedded system for classification respiratory diseases in a real time. *Biomed Signal Process Control* 2024; 90: 105876.

111 Nguyen T, Pernkopf F. Lung sound classification using co-tuning and stochastic normalization. *IEEE Trans Biomed Eng* 2022; 69: 2872–2882.

112 Sharma S, Pandey S, Shah D. Enhancing medical diagnosis with AI: a focus on respiratory disease detection. *Indian J Community Med* 2023; 48: 709–714.

113 Sonali CS, Kiran J, Chinmayi BS, *et al.* Transformer-based network for accurate classification of lung auscultation sounds. *Crit Rev Biomed Eng* 2023; 51: 1–16.

114 Zhang P, Swaminathan A, Uddin AA. Pulmonary disease detection and classification in patient respiratory audio files using long short-term memory neural networks. *Front Med (Lausanne)* 2023; 10: 1269784.

115 Zhang M, Li M, Guo L, *et al.* A low-cost AI-empowered stethoscope and a lightweight model for detecting cardiac and respiratory diseases from lung and heart auscultation sounds. *Sensors* 2023; 23: 2591.

116 Zhang Y, Huang Q, Sun W, *et al.* Research on lung sound classification model based on dual-channel CNN-LSTM algorithm. *Biomed Signal Process Control* 2024; 94: 106257.

117 Zhiqiang S. ICBHI 2017 Challenge. Harvard Dataverse. Harvard, 2023.

118 Palaniappan R, Sundaraj K, Ahamed NU, *et al.* Computer-based respiratory sound analysis: a systematic review. *IETE Tech Rev* 2013; 30: 248–256.

119 Pramono RXA, Bowyer S, Rodriguez-Villegas E. Automatic adventitious respiratory sound analysis: a systematic review. *PLoS One* 2017; 12: e0177926.

120 Moon KS, Lee SQ. A wearable multimodal wireless sensing system for respiratory monitoring and analysis. *Sensors (Basel)* 2023; 23: 6790.

121 Andrès E, Gass R, Charloux A, *et al.* Respiratory sound analysis in the era of evidence-based medicine and the world of medicine 2.0. *J Med Life* 2018; 11: 89–106.

122 Huang D-M, Huang J, Qiao K, *et al.* Deep learning-based lung sound analysis for intelligent stethoscope. *Mil Med Res* 2023; 10: 44.

123 Sfayyih AH, Sulaiman N, Sabry AH. A review on lung disease recognition by acoustic signal analysis with deep learning networks. *J Big Data* 2023; 10: 101.

124 Chen J, Benesty J, Huang Y, *et al.* Fundamentals of noise reduction. *In*: Benesty J, Sondhi MM, Huang YA, eds. Springer Handbook of Speech Processing. Berlin, Heidelberg, Springer Berlin Heidelberg, 2008; pp. 843–872.

125 Alías F, Socoró JC, Sevillano X. A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds. *Appl Sci* 2016; 6: 143.

126 Tu JV. Advantages and disadvantages of using artificial neural networks *versus* logistic regression for predicting medical outcomes. *J Clin Epidemiol* 1996; 49: 1225–1231.

127 Shamshirband S, Fathi M, Dehzangi A, *et al.* A review on deep learning approaches in healthcare systems: taxonomies, challenges, and open issues. *J Biomed Inform* 2021; 113: 103627.

128 Xia T, Han J, Mascolo C. Exploring machine learning for audio-based respiratory condition screening: a concise review of databases, methods, and open issues. *Exp Biol Med (Maywood)* 2022; 247: 2053–2061.

129 Taye MM. Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. *Computers* 2023; 12: 91.

130 Alqudaihi KS, Aslam N, Khan IU, *et al.* Cough sound detection and diagnosis using artificial intelligence techniques: challenges and opportunities. *IEEE Access* 2021; 9: 102327–102344.

131 Jia Z, Chen J, Xu X, *et al.* The importance of resource awareness in artificial intelligence for healthcare. *Nat Mach Intell* 2023; 5: 687–698.

132 Martinez-Martin N, Insel TR, Dagum P, *et al.* Data mining for health: staking out the ethical territory of digital phenotyping. *NPJ Digit Med* 2018; 1: 68.

133 Collins GS, Dhiman P, Andaur Navarro CL, *et al.* Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. *BMJ Open* 2021; 11: e048008.