BMC Pulmonary Medicine

SYSTEMATIC REVIEW Open Access

Inconsistencies in predictive models based on exhaled volatile organic compounds for distinguishing between benign pulmonary nodules and lung cancer: a systematic review

Zhixia Su[1,2], Xiaoping Yu[1], Yuhang He[3], Taining Sha[2], Hong Guo[4], Yujian Tao[5], Liting Liao[6], Yanyan Zhang[7], Guotao Lu[8,9], Guangyu Lu[2*†] and Weijuan Gong[1,6,8*†]

Abstract

Background There is a general rise in incidentally found pulmonary nodules (PNs) requiring follow-up due to increased CT use. Biopsy and repeated CT scan are the most useful methods for distinguishing between benign PNs and lung cancer, while they are either invasive or involves radiation exposure. Therefore, there has been increasing interest in the analysis of exhaled volatile organic compounds (VOCs) to distinguish between benign PNs and lung cancer because it's cheap, noninvasive, efficient, and easy-to-use. However, the exact value of breath analysis in this regard remains unclear.

Methods A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)-oriented systematic search was performed to include studies that established exhaled VOC-based predictive models to distinguish between benign PNs and lung cancer and reported the exact VOCs used. Data regarding study characteristics, performance of the models, which predictors were incorporated, and methodologies for breath collection and analysis were independently extracted by two researchers. The exhaled VOCs incorporated into the predictive models were narratively synthesized, and those compounds that were reported in > 2 studies and reportedly exhibited consistent associations with lung cancer were considered key breath biomarkers. A quality assessment was independently performed by two researchers using both the Newcastle-Ottawa Scale (NOS) and the Prediction Model Risk of Bias Assessment Tool (PROBAST).

Results A total of 11 articles reporting on 46 VOC-based predictive models were included. The majority relied solely on exhaled VOCs ($n = 44$), while two incorporated VOCs, demographical factors, and radiological signs. The variation in the sensitivity, specificity, and AUC indicators of the models that incorporated multiple factors was

†Guangyu Lu and Weijuan Gong contributed equally to this work and should be considered co-corresponding authors.

*Correspondence:
Guangyu Lu
guangyu.lu@yzu.edu.cn
Weijuan Gong
wjgong@yzu.edu.cn

Full list of author information is available at the end of the article

lower compared with those of the models that relied solely on exhaled VOCs. A total of 84 VOCs were incorporated. Of these, 2-butanone, 3-hydroxy-2-butanone, and 2-hydroxyacetaldehyde were identified as key predictors that had significantly higher concentrations in the exhaled breath samples of patients with lung cancer. Substantial heterogeneity was observed in terms of the modeling and validation methods used, as well as the approaches to breath collection and analysis. Many of the reports were missing certain key pieces of clinical and methodological information.

Conclusions Although exhaled VOC-based models for predicting cancer risk might be a conceivable role as monitoring tools for PNs risk, there has been little overall change in the accuracy of these tests over time, and their role in routine clinical practice has not yet been established.

Clinical trial number PROSPERO registration number CRD42023381458.

Keywords Pulmonary nodules, Lung cancer, Volatile organic compounds, Breath biomarkers, Systematic review

Introduction

Lung cancer is the second most common cancer and the leading cause of cancer-related deaths worldwide [1, 2]. Approximately half of patients with lung cancer are diagnosed at an advanced stage, missing the optimal treatment window. This contributes significantly to the high mortality rate associated with the malignancy [3]. Therefore, the early detection and diagnosis of lung cancer are of great significance to both clinical practice and general public health. Chest radiography and sputum cytology have been used for lung cancer screening since the 1970s. However, the sensitivity levels of these modalities are low [4, 5]. Over the past several decades, the US National Lung Cancer Screening Trial (NLST) and several other trials in Europe have demonstrated that low-dose computed tomography (LDCT) scan effectively reduces lung cancer mortality by facilitating earlier-stage diagnoses [6–9]. There is a general rise in incidentally found pulmonary nodules (PNs) requiring follow-up due to increased CT use [10]. However, only a small fraction of these nodules are actually lung cancer [11–14]. Therefore, it is essential to develop efficient and easy-to-use techniques for distinguishing between benign PNs and lung cancer, as this is crucial for guiding clinical decision-making.

Breath analysis, a simple and noninvasive approach, has shown great potential in diagnosing various pulmonary diseases [15]. Breath is a rich medium comprising gas-phase organic and inorganic compounds, as well as aerosols [16]. In the gas phase, there are hundreds of volatile organic compounds (VOCs) of diverse chemical natures that may be present in trace quantities [16]. Although no specific compounds have been identified whose presence or absence in exhaled breath can reliably indicate lung cancer to date, various prediction models based on exhaled VOCs can accurately classify lung cancer patients and healthy volunteers [17–19]. However, the value of breath analysis to distinguish between benign PNs and lung cancer remains unclear.

Published studies on predictive models for distinguishing between benign PNs and lung cancer based on exhaled VOCs have reported conflicting results. First, there is some inconsistency regarding the types and quantities of breath biomarkers incorporated into the models. For example, Peled et al. first developed a diagnostic model based on a single compound (1-octene) to distinguish patients with benign vs. malignant PNs, in their prospective trial [20]. By contrast, a team from the University of Louisville developed PN predictive models by incorporating four compounds: 2-butanone, 3-hydroxy-2-butanone, 2-hydroxyacetaldehyde, and 4-hydroxy-2-hexenal [21–23]. Based on these studies, researchers from Zhejiang University established predictive models by incorporating 19 VOCs to distinguish between benign PNs and lung cancer [24]. It is worth noting that diagnosing lung cancer using a single VOC is challenging, highlighting the importance of incorporating multiple VOCs to conduct more accurate predictions [25, 26]. Second, the performance metrics of existing models have been inconsistent. Two clinical studies have demonstrated that exhaled VOCs can be used to distinguish lung cancers confirmed by biopsy analysis from suspicious PNs observed on repeated LDCT, with high sensitivity and acceptable specificity [23, 27]. However, Liao et al. found that predictive models based solely on exhaled VOCs are not sufficient to accurately identify patients with lung cancer and those with benign PNs and reported that the performance of those models must be improved by combining them with additional factors such as demographic characteristics and radiological findings [28].

This review summarizes the current knowledge regarding exhaled VOC-based predictive models for distinguishing between benign PNs and lung cancer and assesses the overall value of these models in this regard.

Methods

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting checklist [29]. The protocol was registered with the International

Prospective Register of Systematic Reviews (PROSPERO) (registration number CRD42023381458).

Search strategy

A PRISMA-oriented systematic search was performed in PubMed, Embase, Web of Science, Cochrane Library, the Chinese National Knowledge Infrastructure (CNKI), China Biology Medicine disc (CBM), Wanfang, and CQVIP. We initially searched databases from their inceptions until 23 July 2022, with an update on 30 September 2023. The search was conducted through the combination of the Medical Subject Headings (MeSH) terms and keywords 'lung neoplasms' 'lung cancer' 'multiple pulmonary nodules' 'solitary pulmonary nodule' 'volatile organic compounds' 'breath' and 'exhaled'. A list of the detailed search strategy used is described in the Appendix 1.

Inclusion and exclusion criteria

We included studies that reported on the development and validation of exhaled VOC-based risk-predictive models to distinguish between benign PNs and lung cancer. Studies were included if they were conducted in patients with LC confirmed via biopsy analysis, as well as in patients with benign PNs confirmed via biopsy analysis or detected by repeated radiological scans [30]. The included studies needed to identify and report specifically which VOC biomarkers were used. The predictors incorporated into the models were permitted to have been either exhaled VOCs alone or combined with other factors such as demographic characteristics and radiological signs. Articles published in English or Chinese were included, without any restrictions on study design.

Studies that analyzed exogenous VOCs or compounds in breath condensate and biofluids such as serum, urine, feces, and gastric content were excluded. Studies were also excluded if they were not carried out in humans or were not relevant to distinguishing between benign PNs and lung cancer. Studies limited to identifying risk predictors, reviews, letters, comments, and conference abstracts were also excluded.

Study selection

The electronic reports identified were imported into the reference manager Endnote and duplicates removed. Screening title and abstract of studies identified was performed by two researchers (ZXS and GYL) independently, and then full texts were reviewed to determine eligibility for inclusion. During this procedure, potential studies from the reference lists of original articles on this issue were screened and reviewed thoroughly for eligibility.

Data extraction and analysis

A pre-defined table was designed to extract the variables through a panel discussion with experts and epidemiologists. Information was extracted from each study includes: (1) the characteristics of study (title, first author, publication year, study country, study design, and study participants); (2) the characteristics of prediction models (candidate variables of the models, variables incorporated into the models, statistical method, modeling method, internal validation, external validation, sensitivity, specificity, accuracy, and area under the receiver operating characteristics (ROC) curves); (3) detailed methodologies (breath test environments, patient physiological conditions, sample collection methods, and VOC analysis and identification). Data extraction was independently performed by two researchers (ZXS and GYL). When encountering disagreements, our research team would discuss the article and reach a consistent agreement.

Quality assessment

The adapted Newcastle-Ottawa Scale (aNOS) used as a tool for risk of bias assessment of the included articles (Appendix 2) [31]. Moreover, the models' applicability to the intended population and setting were assessed by the Prediction Model Risk of Bias Assessment Tool (PROBAST) (Appendix 3) [32, 33]. Two researchers conducted a critical appraisal of the studies, with any disagreements re-solved through consensus.

Results

Characteristics of the included studies

A total of 2,288 studies were identified. Among them, 2,240 were published in English, while 48 were in Chinese. In addition, five additional studies have been preliminary added according to their titles through other sources, including reference and website [28, 34–37]. After removing 1303 duplicates, a total of 990 studies underwent title and abstract screening, before 450 studies were screened for full-text. A total of 439 studies were excluded for various reasons, as outlined in Appendix 8. Our final analysis included 11 studies (Fig. 1 and Appendix 7). These studies reported 46 VOC-based predictive models that had been used to distinguish a total of 597 patients with benign PNs from 1,700 patients with lung cancer. The number of participants involved in establishing the predictive models ranged between 72 and 768 per study. Of the 11 total studies, seven (63.64%) were conducted in the US and four (36.36%) were carried out in China. Nine (81.82%) were cross-sectional studies, one was a prospective trial, and one was a cohort study. The characteristics of the included studies are listed in Table 1.

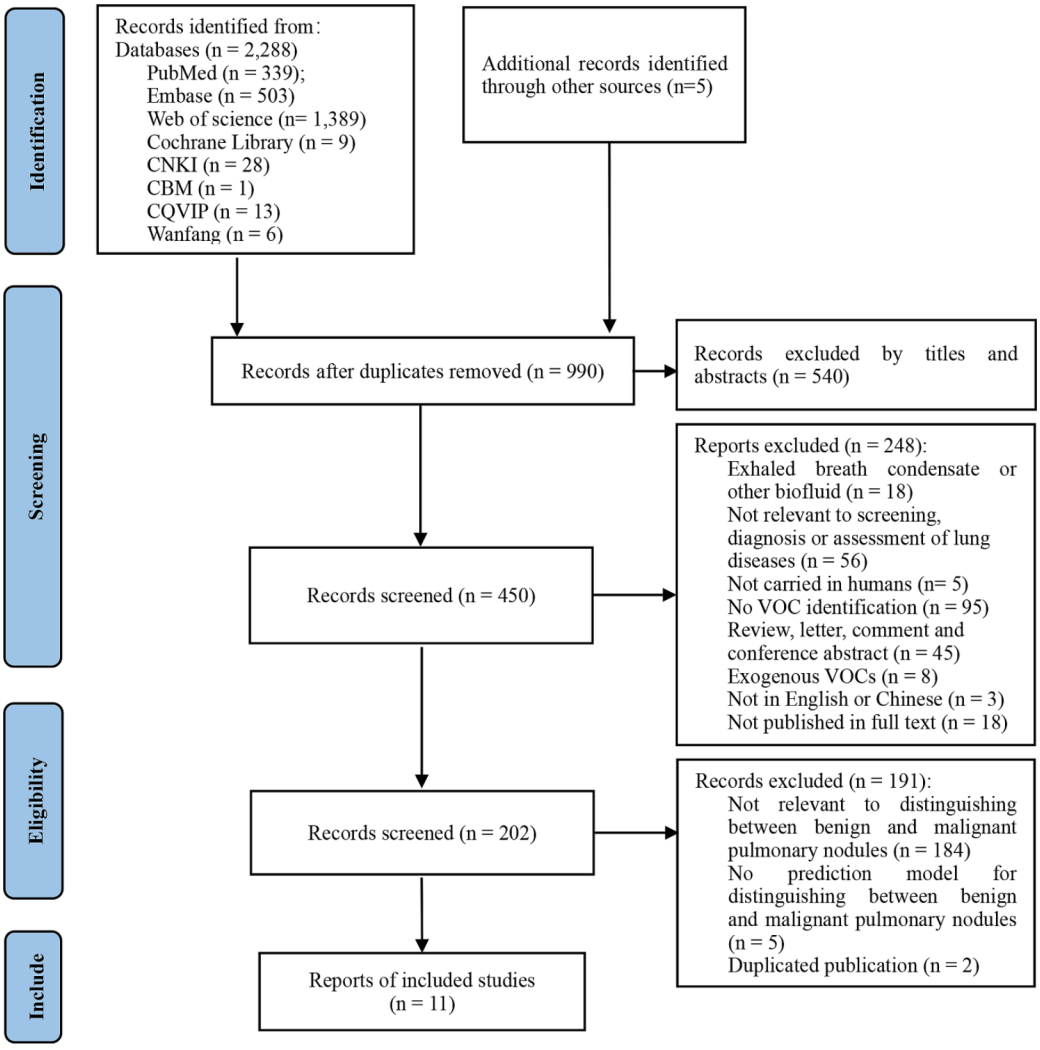

Fig. 1 Selection process of PRISMA flow diagram

Breath collection and analysis methods

The most commonly used methodology for collecting exhaled breath involved the use of Tedlar bags (n=8 studies, 72.73%), although the sample volume ranged between 450 and 1,000 mL. All of the studies used MS-based techniques such as gas chromatography-mass spectrometry (GC-MS; n=5, 45.45%), Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS; n=5, 45.45%), and high-pressure photon ionization time-of-flight mass spectrometry (HPPI-TOFMS; n=1, 9.1%). Notably, solid-phase microextraction (SPME; n=3, 27.27%) and thermal desorption (TD; n=2, 18.18%) were the most frequently used methods for pre-concentrating the samples. Regarding the identification of the chemical structures of the VOCs, four of the studies (36.36%) referred to the National Institute of Standards and Technology Library (NIST), while two (18.18%) relied on MS techniques and retention times.

The factors that influenced exhaled VOC concentrations were divided into three categories: breath test environments, patient physiological conditions, and breath collection and analysis methods. A total of seven studies (63.64%) described their breath test environments, while seven (63.64%) focused on the physiological conditions of the patients analyzed. All of the included studies discussed their breath collection and analysis methods. Detailed information regarding these factors is presented in Table 2.

Incorporated predictors

The number of predictors incorporated into the models ranged between 1 and 35 (Table 3). Out of 46 total predictive models, 44 (95.64%) were solely based on VOC biomarkers. Only one model (2.18%) incorporated both VOC biomarkers and patient ages. Another model (2.18%) considered VOC biomarkers, age, and

Table 1 Characteristics of the included prediction model studies

Study	Year	Country	Study design	Sample size	Study participants	BPN [*n* (%)]	Patient characteristic information		
							Setting	Patients characteristics	Diagnostic methods
Peled, N. [20]	2012	America	prospective trial	72	53 patients with MPN, 19 with BPN	19(26.39)	the University of Colorado Cancer Center and Denver Veterans Affairs Medical Center	47 NSCLC (30 adenocarcinomas, 13 squamous cell carcinoma, 2 large cell carcinoma, and 2 poorly differentiated carcinoma), 6 SCLC	Histologic diagnosis or serial CT
Bousamra, M., 2nd [21]	2014	America	cross-sectional	147	107 patients with LC, 40 patients with BPNs	40(27.21)	NR	97 NSCLC, 8 SCLC, 1 combined SCLC and NSCLC, 1 carcinoid	Bronchoscopy, CT-guided biopsy, or surgical biopsy/ serial CT scans
Fu, X. A. [22]	2014	America	cross-sectional	129	97 patients with MPN, 32 patients with BPN	32(24.80)	NR	88 NSCLC (33 adenocarcinoma, 32 squamous cell carcinoma), 9 SCLC	Biopsy or resection
Li, M. X. [61]	2015	America	cross-sectional	119	85 patients with MPN, 34 patients with BPN	34(28.57)	the James Graham Brown Cancer Center at the University of Louisville	78 NSCLC (31 adenocarcinoma, 33 squamous cell,14 Others), 7 SCLC;2 pneumonia, 8 granuloma, 3 inflammation, 2 histoplasmosis, 2 epithelial cells, 1 foamy macrophages, 16 unknown	CT, pathological diagnoses
Schumer, E. M. [23]	2015	America	cross-sectional	221	156 patients with LC, 65 patients with BPNs	65(29.41)	NR	103 stage 0-II, 53 stage III-IV	Pathologic/tissue diagnosis or repeated CT
Phillips, M. [27]	2019	America	cross-sectional	80	65 patients with MPN, 15 patients with BPN	15(18.75)	NR	NR	LDCT, Biopsy
Shaohua Xie [49]	2020	China	cross-sectional	147	104 with MPN, 43 patients with BPN	43(29.25)	Sichuan Cancer Hospital	NR	CT, Pathologic diagnosis
Pengqiang Liao [28]	2021	China	cross-sectional	768	629 patients with LC, 139 patients with BPN	139(18.10)	Sichuan Cancer Hospital; Nanchong Central Hospital; Longquanyi CDC	NR	Surgery or biopsy
Chen, X. [24]	2021	China	cross-sectional	230	160 patients with MPN, 70 patients with BPN	70(30.43)	Sir Run Run Shaw Hospital	149 NSCLC (117 adenocarcinomas,32 squamous cell carcinoma), 10 SCLC	CT, X-ray, pathological diagnoses
Rai, S. N. [58]	2022	America	cross-sectional	221	156 patients with LC, 65 patients with benign PNs	65(29.41)	healthy control: from patient family members the James Graham Brown Cancer Center and Jewish Hospital at the University of Louisville	103 stage 0-II, 53 stage III-IV	CT, PET scans; biopsy or surgically resected specimens
Ding, X. [60]	2023	China	cohort study	PKUPH Cohort: 100 CHPKU Cohort: 63	PKUPH Cohort: 49 lung cancer patients and 51 benign pulmonary nodules CHPKU Cohort: 39 lung cancer and 24 benign nodules	PKUPH Cohort: 51(51.00) CHPKU Cohort: 24(38.10)	PKUPH Cohort: the Department of Thoracic Surgery, Peking University People's Hospital; CHPKU Cohort: the Cancer Hospital of Peking University	PKUPH Cohort: stage I; 49 adenocarcinom CHPKU Cohort: stage I-III; 19 adenocarcinom, 2 quamous cell carcinoma, 2 small cell lung cancer	CT, pathological diagnoses

NR: Not report; MPN: Malignant pulmonary nodules; BPN: Benign pulmonary nodules; LC: Lung cancer; NSCLC: Non-small cell lung cancer; SCLC: Small cell lung cancer; CT: Computerized tomography; PET: positron emission tomography; PKUPH Cohort: Peking University People's Hospital as the discovery cohort; CHPKU Cohort: the Cancer Hospital of Peking University

radiological signs (including nodule size, count, type, and spiculation).

A total of 84 VOCs were incorporated into the predictive models. The most commonly reported VOCs were 2-hydroxyacetaldehyde ($n=5$ studies), 2-butanone ($n=4$), 3-hydroxy-2-butanone ($n=4$), and butyric acid ($n=3$). Among these VOCs, the concentrations of 2-hydroxyacetaldehyde, 2-butanone, and 3-hydroxy-2-butanone were reported to have been significantly higher in the exhaled breath samples of patients with lung cancer. The associations between butyric acid concentration and lung cancer were not reported among any of the included studies. Table 4 presents a comprehensive summary of all of the VOCs that were reported on in ≥2 of the studies. The VOCs that were reported on in only one study are listed in Appendix 4, along with their chemical classes and CAS registry numbers.

Modeling methods

Table 3 presents detailed information regarding the modeling methods used in the included studies. The most commonly used modeling methods were machine learning—which included support vector machine classification models ($n=9$, 19.57%), logistic regression analysis ($n=6$, 13.04%), random forest algorithm ($n=3$, 6.52%), discriminant analysis ($n=3$, 6.52%), artificial neural network ($n=2$, 4.35%), and partial least squares analysis ($n=1$, 2.17%). The statistical modeling methods used for 22 of the models (47.83%) were unclear.

Predictive performance

Each study included in our analysis reported at least one measure of predictive performance (Table 3). Nine of the studies reported classification measures of 20 predictive models. Of these, the sensitivity and specificity values of nine of the models (45%) were reported to both be >70%. Six of the studies reported index of prediction accuracy values that ranged between 54.15% and 94.6%. Eight of the studies assessed discrimination using AUC values, which ranged between 0.625 and 0.986. Notably, none of the studies reported on calibration measures.

The variation in terms of sensitivity, specificity, and AUC values of the models that incorporated multiple factors was lower compared with those of the models that were based solely on exhaled VOCs. Among the models based solely on exhaled VOC biomarkers, the model with the highest sensitivity (100%) exhibited a specificity of 81.8%—while the model with the highest specificity (100%) showed a sensitivity of 28%. When exhaled VOC biomarkers were combined with other factors, the model with the highest sensitivity (80.80%) had a specificity of 60.50%, while the model with the highest specificity (68.3%) had a sensitivity of 78.7%. The models based solely on exhaled VOCs had AUC values ranging between 0.625 and 0.986, while those that incorporated additional factors such as age and radiological signs maintained stable AUC values of 0.776–0.781.

Validation approaches

Out of the 46 developed models, 19 (41.3%) exclusively underwent internal validation using methods such as K-fold cross-validation ($n=13$, 28.3%), leave-group-out cross-validation (LGOCV; $n=5$, 10.9%), and leave-one-out cross-validation (LOOCV; $n=1$, 2.2%). Two of the models (4.3%) were internally validated through K-fold cross-validation before being subjected to blinded validation at two independent laboratories. Only one study validated two models (4.3%) using a new patient sample.

Study quality assessment

The aNOS scores of the included studies ranged between 6 and 9 (Appendix 5). Of the 11 total studies, five were rated very good, five were rated good, and one was rated satisfactory. The one study that received satisfactory ratings did not report on the representativeness levels of their patient samples.

The PROBAST tool was used to evaluate the risk of bias associated with the included studies. As shown in Fig. 2 and Appendix 6, all were found to have a high risk of bias in terms of participants, predictors, analysis, and overall domains. This high risk of bias may be attributable to the inappropriate inclusion criteria or small sample sizes of the studies. Predictors were also evaluated based on outcome information, and the selection of predictors using univariable analysis was done without reporting on measures of discrimination or calibration for model performance. Furthermore, the risk of bias in the outcome domains was unclear for all 11 of the included studies.

In terms of the applicability of the models, nine of the studies had low risks of bias in the participant domain, whereas two had unclear risks of bias. All of the models developed had unclear risks of bias in the predictors, outcomes, and overall domains.

Discussion

This systematic review assessed the performance of 46 predictive models for lung cancer, based on VOCs detected in exhaled breath samples, reported in 11 studies and involving a total of 2,297 patients. Although the findings indicated that there was significant heterogeneity in the predictive performances of the models in terms of distinguishing between benign PNs and lung cancer, the models that incorporated additional factors such as demographic characteristics and radiological signs showed better performance metrics. Several compounds—including 2-butanone, 3-hydroxy-2-butanone, and 2-hydroxyacetaldehyde—were identified as being the most significant VOCs for distinguishing between

benign PNs and lung cancer. Despite the potential of VOC breath analysis, there has been little overall change in the accuracy of these tests over time, and their role in routine clinical practice requires further research to be established.

Although the concentrations of exhaled 2-butanone, 3-hydroxy-2-butanone, and 2-hydroxyacetaldehyde have been identified as significant biomarkers for distinguishing between benign pulmonary nodules (PNs) and lung cancer, the endogenous origins of these volatile organic compounds (VOCs) remain largely unknown or speculative. In lung cancer patients, elevated levels of ketones have been observed, which may result from fatty acid oxidation known to increase the production of 2-butanone [38, 39]. A previous study has demonstrated that exhaled 3-hydroxy-2-butanone is likely an endogenous product resulting from the degradation of 2-butanone [40]. Similar research in rodents has shown that 2-butanone is oxidatively metabolized to 3-hydroxy-2-butanone, presumably by cytochrome P-450-dependent monooxygenases [41, 42]. Additionally, 3-hydroxy-2-butanone is produced during the detoxification of acetaldehyde and may participate in pulmonary redox cycling, potentially generating toxic reactive oxygen species (ROS) that can damage lung tissue [43–45]. Endogenous 2-hydroxyacetaldehyde is formed through the oxidative degradation of glucose, as well as from glycated proteins, lipid peroxidation, and the oxidation of amino acids [46]. This compound, produced by human neutrophils during phagocytosis, can be a potential source of ROS and may play an important role in tumor development and progression [46]. It is also important to note that exhaled isoprene has been identified as a marker for lung cancers [47]. Initially, it was believed to originate from hepatic cholesterogenesis; however, recent study has revealed that exhaled isoprene originates from muscular lipolytic cholesterol metabolism, as determined by the IDI2 gene [48]. Therefore, further research is needed to uncover the human metabolic origins of endogenous VOCs in lung cancer to promote clinical validation for diagnosis.

Our literature review indicated that, in addition to VOCs alone, incorporating multiple factors, such as patient characteristics and radiological signs, can greatly improve the stability of predictive models based on breath samples. Although the majority of the studies included in this review established predictive models based solely on exhaled VOCs, two of the ones that exhibited superior levels of performance also incorporated patient ages and radiological signs [28, 49]. When considering additional factors, variables such as age and smoking status should be considered first, as these represent known predictors of lung cancer risk [50, 51]. Furthermore, researchers have shown that diagnosing lung cancer using a single VOC is challenging and that incorporating several VOCs may be necessary to enhance the performance of such predictive models [25, 26]. However, including too many variables may increase the risk of model overfitting and spurious relationships [52]. Without proper statistical corrections, spurious correlations can be found even in entirely nonsensical contexts, referred to as Voodoo correlations [53]. This issue is further exacerbated by a relatively small sample size, which reduces the power of statistical tests and increases the likelihood of false positives. Therefore, well-established techniques such as the Bonferroni correction or False Discovery Rate (FDR) control that can and should be used for the reducing the likelihood of spurious correlations [54, 55].

Performance of predictive models varied significantly depending on the algorithm of machine learning. The main application field of logistic regression (LR) and support vector machine (SVM) is binary classification which makes these algorithms attractive to solve these clinical tasks [56, 57]. Xie SH, Liao PQ, and Chen X et al. utilized LR algorithm to develop predictive models based on exhaled VOCs to distinguish between benign PNs and lung cancer [24, 28, 49]. The predictive performance was almost the same for these models, with acceptable AUC values. Rai et al. trained SVM to establish their relevance in lung cancer patients' classification which also achieved an acceptable accuracy [58]. The Random forest (RF) algorithm is more flexible because it can identify a broader scope of possible relationships between the model predictors and the disease status [59]. For example, Ding, X, et al. demonstrated the efficacy of the RF models based 16 exhaled VOCs for discriminating lung cancer from benign PNs [60]. In addition, validation approaches also varied widely between studies, including K-fold cross-validation [24, 28, 58] leave-group-out cross-validation [61], and leave-one-out cross-validation [20], which makes replication of results between studies difficult.

Our review further highlighted that sample collection and measurement techniques are hard to rationalize in clinical perspective. Many researchers have optimized method of exhaled breath analysis before the study of real subjects, but the results are varied. For example, some studies have required participants to fast for 6–12 h before breath collection [24, 28, 49], while Li MX. et al. conducted in patients without any diet controls [61]. In addition, some researchers have utilized Tedlar bags, Mylar bags [20–23, 61], and Bio-VOC breath samplers [49, 62] to collect breath samples. While others preferred portable breath sample collection devices or self-made collection devices [63, 64]. Only the end-tidal phase of a breath represents systemic concentrations of VOCs [65], while the lack of alveolar sampling in the included studies is biased via various confounding effects [21–23, 58, 60,

Table 2 Summary of factors reported to influence volatile organic compounds within exhaled breath

Study	Environmental considerations			Patient-related factors		Breath collection methods							Method of analysis			
	Ambient air	Specific room for breath sampling	Detection contamination	Physiological conditions	Setting	Sampling period	Sample collection and extraction methods	Breath portion	Volume	Route	Flow rate	Stability of VOCs	Analysis platform	Sensitivity	Repeatability	VOC identification
Peled, N. [20]	NR	NR	NR	NR	the University of Colorado Cancer Center and Denver Veterans Affairs Medical Center, America	NR	Mylar bag + SPME	all	750 mL	mouth	NR	NR	GC-MS	NR	NR	NR
Bousamra, M., 2nd [21]	collected.	NR	NR	NR	NR	NR	Teldar bag	all	1 L	NR	NR	NR	FT-ICR-MS	NR	NR	NR
Fu, X. A. [22]	collected.	NR	NR	NR	NR	before resection	1 L Tedlar bag	all	1 L	NR	NR	NR	FT-ICR-MS	NR	NR	confirmed by FT-ICR-MS/MS
Li, M. X. [61]	collected.	clinic room	Tedlars bags and syringes were tested free of contamination.	no diet controls	the James Graham Brown Cancer Center at the University of Louisville, America	NR	1 L Tedlar bag	all	NR	NR	5 mL/min	NR	FT-ICR-MS	NR	NR	NR
Schumer, E. M. [23]	NR	NR	NR	NR	NR	NR	Tedlar bag	all	1 L	NR	NR	NR	FT-ICR-MS	NR	NR	NR
Phillips, M. [27]	NR	NR	NR	NR	NR	NR	a dual-bedded sorbent trap + TD	all	1 L	mouth	NR	at least 3 years	GC-MS	NR	replicate assays at two independent laboratories	according to mass spectrometry library NIST

Table 2 (continued)

Study	Environmental considerations			Patient-related factors		Breath collection methods							Method of analysis			
	Ambient air	Specific room for breath sampling	Detection contamination	Physiological conditions	Setting	Sampling period	Sample collection and extraction methods	Breath portion	Volume	Route	Flow rate	Stability of VOCs	Analysis platform	Sensitivity	Repeatability	VOC identification
Shaohua Xie [49]	collected.	a ventilated room	NR	Sit over 10 min, avoid food and smoke for at least 8 h. No deep inhalation before sampling, no ventilation through nasal and no second breaths during sampling.	Sichuan Cancer Hospital, China	before biopsy or resection	500 mL Tedlar bag + Bio-VOC breath sampler + SPME	last portion	450mL	mouth	NR	1 week	GC-MS	35–200 m/z	NR	according to mass spectrometry library NIST 08
Pengqiang Liao [28]	NR	NR	NR	Participants were required not to eat for 8 h before breath collection, and rest for at least 10 min in a well-ventilated room; Deep breathing before sampling and nasal ventilation and secondary breathing during sampling were not allowed.	Sichuan Cancer Hospital; Nanchong Central Hospital; Longquanyi CDC	before resection	500 ml Tedlar bag + SPME	last portion	450mL	mouth	NR	within 7 days	GC-MS	35-200amu	NR	according to mass spectrometry library NIST and RI

Table 2 (continued)

Study	Environmental considerations			Patient-related factors		Breath collection methods							Method of analysis			
	Ambient air	Specific room for breath sampling	Detection contamination	Physiological conditions	Setting	Sampling period	Sample collection and extraction methods	Breath portion	Volume	Route	Flow rate	Stability of VOCs	Analysis platform	Sensitivity	Repeatability	VOC identification
Chen, X. [24]	collected.	a specific controlled room	NR	Avoid high fat food any antioxidant supplements for at least 24 h. Fasting for 12 h, then stayed in a specific room. Next day before the breakfast, the participants arrived in a specific controlled room for breath sample collection.The participants didn't brush teeth, and gargled with plain tap water 15 min before.	Sir Run Run Shaw Hospital, China	NR	TD	last portion	1 L	mouth	NR	NR	GC-MS	45–500 m/z	Repetitive tests were performed on collected air and breath samples.	according to mass spectrometry library NIST 05 and NIST 05 s

Table 2 (continued)

Study	Environmental considerations			Patient-related factors		Breath collection methods							Method of analysis			
	Ambient air	Specific room for breath sampling	Detection contamination	Physiological conditions	Setting	Sampling period	Sample collection and extraction methods	Breath portion	Volume	Route	Flow rate	Stability of VOCs	Analysis platform	Sensitivity	Repeatability	VOC identification
Rai, S. N. [58]	collected.	a clinic exam room	NR	NR	healthy control: patient family members the James Graham Brown Cancer Center and Jewish Hospital at the University of Louisville	NR	1 L Tedlar bag	all	1 L	mouth	5 mL/min	NR	FT-ICR-MS	NR	NR	NR
Ding, X. [60]	collected.	a fixed room	NR	Participants fasted for at least 6 h. Participants were asked not to ingest spicy food, alcohol, or coffee the night before exhaled breath collection. Participants first gargled with pure water	PKUPH Cohort: the Department of Thoracic Surgery, Peking University People's Hospital; CHPKU Cohort: the Cancer Hospital of Peking University	before PET-CT scanning and the morning before surgery	Tedlar bag	all	NR	mouth	NR	within 4 h	HPPI-TOFMS	m/z < 500	NR	detected by HPPI-TOFMS

NR: Not report; GC-MS: Gas Chromatography Mass Spectrometry; LC: Lung cancer; SPME: Solid phase microextraction; NIST: the National Institute of Standards and Technology Mass Spectrometry library; FT-ICR-MS: Fourier Transform Ion Cyclotron Resonance Mass Spectrometry; TD: thermal desorption; PTR-MS: Proton Transfer Reaction Mass Spectrometry

Table 3 Predictors, methods, and performance of the assessed development models

Study	Study participants	Included variables		Statistical method	Modeling method	Validation method	Number	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC
		VOCs	no-VOCs								
Peled, N. [20]	53 patients with MPN, 19 with BPN	1-octene	NR	Wilcoxon/Kruskal– Wallis test	DFA	Internal validation: LOOCV	1	① 86%	① 96.00%	① 88.00%	① 0.986
Bousamra, M., 2nd [21]	107 patients with LC, 40 patients with BPNs	2-butanone; 3-hydroxy-2-butanone; 2-hydroxyacetaldehyde; 4-hydroxyhexenal	NR	Wilcoxon test + ANOVA	NR	NR	9	① NR; ② 21%; ③ 60%; ④ 84%; ⑤ 92%; ⑥ 28%; ⑦ 70%;⑧ 89%; ⑨ 95%	① NR; ② 100%; ③ 94%; ④ 74%; ⑤ 45%; ⑥ 100%; ⑦ 95%; ⑧ 74%; ⑨ 45%	① -⑨ NR	① 0.8; ② -⑨ NR
Fu, X. A. [22]	97 patients with MPN, 32 patients with BPN	2-butanone; 2-hydroxyacetaldehyde; 3-hydroxy-2-butanone; 4-hydroxyhexenal	NR	Wilcoxon test	NR	NR	1	① 89.80%	① 81.30%	① 87.60%	① NR
Li, M. X. [61]	85 patients with MPN, 34 patients with BPN	hydroxyacetaldehyde; 2-butanone; 4-hydroxy-2-hexenal; a mixture of 2-pentanone and pentanal; 3-hydroxy-2-butanone	NR	Kruskal– Wallis test	PLS + SVM + RF + LDA + QDA	Internal validation: LCOGV	14	① 100%; ② -⑭ NR	① 81.8%; ② -⑭ NR	① 0.946; ② 0.892; ③ 0.892; ④ 0.865; ⑤ 0.811; ⑥ -⑭ NR	① -⑤ NR;⑥ 0.901; ⑦ 0.821; ⑧ 0.694; ⑨ 0.793; ⑩ 0.625; ⑪ 0.696; ⑫ 0.793;⑬ 0.66; ⑭ 0.874
Schumer, E. M. [23]	156 patientd with LC, 65 patients with BPNs	2-butanone, 3-hydroxy-2-butanone, 2-hydroxyacetaldehyde, 4-hydroxyhexanal	NR	NR	NR	NR	3	① 93.6%; ② 76.9%; ③ 40.4%	① 44.6%; ② 78.5%; ③ 90.8%	① -③ NR	① -③ NR
Phillips, M. [27]	65 patients with MPN, 15 patients with BPN	1,4-butanediol; 2-pentanamine, 4-methyl-; 2-propanamine; 3-butenamide; 4-penten-2-ol; acetamide, 2-cyanoalanine; N-methylglycine; octodrine	NR	NR	fuzzy logistic regression	Blinded validation; Internal validation: cross-validation	2	① 80.1%; ② NR	① 75%; ② NR	① 80.5%; ② NR	① 0.88; ② 0.80
Shaohua Xie [49]	104 with MPN, 43 patients with BPN	Cyclopentane; Pentane, 3-methyl-; ethylbenzene; N, N-di(methyl)formamide	age	t-test + Mann-Whitney U test + χ 2 test + Fisher exact probability	bivariate logistic regression	NR	1	① 80.80%	① 60.50%	① NR	① 0.781

Table 3 (continued)

Study	Study participants	Included variables		Statistical method	Modeling method	Validation method	Number	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC
		VOCs	no-VOCs								
Pengqiang Liao [28]	629 patients with LC, 139 patients with BPN	α -Pinene; 2,3,6-trimethyl-Heptane; 2-Methylnaphthalene; Tetrachloroethylene; Naphthalene; Naphthalene, 1-methyl-; Furan, tetrahydro-2,2,5,5-tetramethyl; p-Cresol; 1-Propene-1-thiol; 5-Hepten-2-one,6-methyl-; Undecane; Methylvinylketone; Dimethadione; 3-Pentenoicacid,4-methyl; 4-Methoxyphenol; Propanoicacid,2-methyl-,butylester; N,N-Dimethylacetamide; tert-Butanol; Methylenechloride; Guaiacol; AceticAcid; benzene; PropionicAcid; Oxetane,2,2-dimethyl; Acetone; Butanoicacid; Cyclohexane and 3 unknown compounds	age, emphysema, tumor size, tumor type, spicule sign	—	GA-SVM + Binary logistic regression	Internal validation:5-fold cross-validation	2	① 51.06%; ② 78.7%	①68.29%; ②68.3%	①54.15% ②75.56%	①0.65; ②0.776
Chen, X. [24]	160 patients with MPN, 70 patients with BPN	1,2-Dichloroethane; Benzene; 2,4-Dimethylhexane; Heptane; Di-tert-butylperoxide; Tetrachloroethylene; Heptanal; Propylcyclohexane;2-Nonenal, (2E); Cyclohexane, heptyl; 1,4-Methanoazulen-9-ol, decahydro-1,5,5,8a-tetramethyl-,(1R,3aR,4 S,8aS,9 S);2,4-Hexadiyne; Butyricacid; 2-methylbutyricacid; Cyclononasiloxane, octadecamethyl-(8CI,9CI); Octadecamethylcyclononasiloxane; Benzoicacid,3,5-bis(1,1-dimethylethyl); METHYL10-METHYLUNDECANOATE; p-Terphenyl	NR	NR	logistic regression + ANN	Internal validation:5-fold cross-validation	4	①-④ NR	①-④ NR	①-④ NR	① 0.809; ② 0.799;③0.756; ④ 0.779;
Rai, S. N. [58]	156 patients with LC, 65 patients with benign PNs	Butyraldehyde; Hexanal;2-Heptanone; Octanal; Undecanal; Butyricacid; Aceticacid; Acrolein; 4-hydroxyhexanal; 4-Hydroxynonenal; Cyclopentanone; Dicyclohexylketone	NR	Boot-SVM-RFE	SVM	Internal validation: 5-fold cross-validation	7	①-⑦NR	①-⑦NR	① 76.98;② 76.07;③ 76.25;④ 78.48;⑤ 76.72; ⑥ 77.65;⑦ 77.92	① -⑦NR
Ding, X. [60]	PKUPH Cohort: 49 lung cancer patients and 51 benign pulmonary nodules CHPKU Cohort: 39 lung cancer and 24 benign nodules	acetaldehyde, 2-hydroxyacetaldehyde, isoprene, pentanal, butyricacid, toluene,2,5-dimethylfuran, cyclohexanone, hexanal, heptanal, acetophenone, propylcyclohexane, octanal, nonanal, decanal, and2, 2-dimethyldecane	NR	t-test + Fisher's exact test	RF	External validation	2	① 82.1%; ② 71.8%	① 92.3%;② 76.9%	① 84.6%;② 73.1%	①0.872; ②0.744

①-⑭: model 1 to 14; NR: Not report; Forward SDA: Forward stepwise discriminant analysis; LOOCV: Leave-one-out cross-validation; Fisher SDA: Fisher stepwise discriminant analysis; DFA: Discriminant factor analysis; ANOVA: Analysis of variance; ANN: Artificial neural network; Fisher DA: Fisher discriminant analysis; GA-SVM: Genetic algorithm-support vector machine; PLS: Partial least squares; SVM: Support vector machine classification models; RF: Random forest algorithm; LDA: Linear discriminant analysis; QDA: Quadratic discriminant analysis; LGOCV: Leave group out cross-validation

Table 4 Volatile organic compounds reported in at least two studies

Compound name	Frequency	Level of VOCs		Chemical Classes	CAS
		LC> BPN	BPN> LC		
2-hydroxyacetaldehyde	5	Yes		Aldehydes	141-46-8
4-hydroxyhexenal	2	Yes		Aldehydes	109710-37-4
Acrolein	2	Yes		Aldehydes	107-02-8
4-hydroxy-2-nonenal	2	Yes		Aldehydes	75899-68-2
Acetaldehyde	2	NR	NR	Aldehydes	141-46-8
4-hydroxyhexanal	2	Yes		Aldehydes	109710-36-3
Heptanal	2	NR	NR	Aldehydes	111-71-7
Hexanal	2	Yes		Aldehydes	66-25-1
Octanal	2	Yes		Aldehydes	124-13-0
2-butanone	4	Yes		Ketones	78-93-3
3-hydroxy-2-butanone	4	Yes		Ketones	513-86-0
Butyric acid	3	NR	NR	Acids	107-92-6
Acetic Acid	2	NR	NR	Acids	64-19-7
Propylcyclohexane	2	NR	NR	Alkanes	1678-92-8
Benzene	2	NR	NR	Aromatic compounds	71-43-2
Tetrachloroethylene	2	NR	NR	Haloalkanes	127-18-4

NR: Not Reported; CAS: Chemical abstracts service; LC: lung cancer; BPN: benign pulmonary nodules

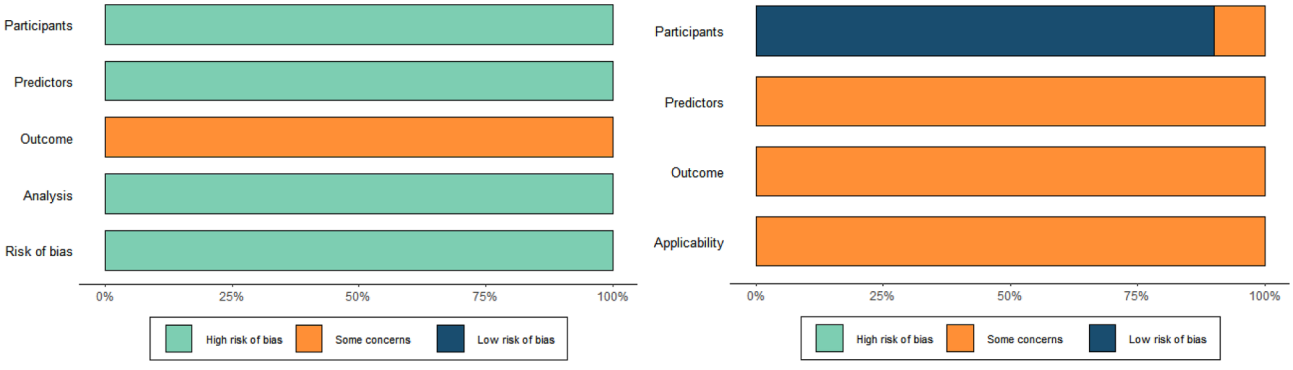

Fig. 2 Risk of bias and applicability assessment according to the PROBAST

61]. Physio-metabolic and analytical confounders are to be minimized to realize actual pathophysiological effects on exhaled VOC concentrations [66]. Therefore, to standardize the procedure for exhaled breath analysis, expert opinions can be gathered through various methods, such as the Delphi process [67].

MS-based techniques are generally considered the gold standard for analyzing VOCs, owing to their ability to determine the molecular masses and possible chemical structures of individual VOCs [68]. However, this technique is subject to certain limitations. First, the analysis of VOC biomarkers using MS-based techniques requires pre-concentration and a high level of expertise, making them too expensive and complex for many clinical application [69, 70]. Moreover, pre-concentration methods may selectively enhance the signals of certain VOCs while simultaneously leading to the loss of others [20]. Second, breath samples are recommended to be analyzed within six hours of collection to best preserve sample compositions [71]. However, it is important to consider that background pollutants from the sampling equipment may potentially alter breath sample compositions as well [71, 72]. From a clinical perspective, clinicians are more interested in identifying "treatable traits" by stratifying patients based on clinically relevant qualities such as diagnosis, prognosis, and treatment response [73, 74]. In this scenario, simpler more cost-effective techniques have been considered and developed, such as the electronic nose (or "e-nose") [75]. Many researchers have concluded that e-noses have the potential to become promising diagnostic tools in everyday clinical practices [76, 77].

Although existing VOC-based predictive models have shown acceptable levels of performance for distinguishing between benign PNs and lung cancer, the majority of the existing predictive models for distinguishing between benign PNs and lung cancer have not undergone validation, which hinders their clinical implementation. Ideally, external validations of VOC-based predictive models should be performed in large observational cohorts that have been carefully designed to be accurately

representative of real-world patients with benign PNs and lung cancer [78]. To better inform clinical practices, future studies should carefully consider the heterogeneity of prediction effects and conduct model validations to develop predictive models that are of value to clinicians.

Limitations

This review summarized the current knowledge on exhaled VOC-based predictive models for distinguishing between benign PNs and lung cancer. However, it was subject to certain limitations worth noting. First, there was substantial heterogeneity among the included studies in terms of the methods used for their establishment, as well as for patient breath sample collection and analysis. Consequently, no quantitative meta-analysis has been conducted on the results of the studies. Second, all of the included studies were found to have a high risk of bias, likely caused by inappropriate inclusion criteria, the selection of predictors using univariate analyses, or a lack of indicators that could be used to evaluate calibration risk in the models. Third, most included studies in this review were conducted with cross-sectional study design, which conducted in patients with mid-/late-stage LC. The difference in performance of models based on breath VOCs among benign PNs and early LC demand further investigation with a large sample. In addition, eight of the studies were conducted in small groups, which may have limited the internal validity of their methods and the generalizability of their models to the general population. Finally, studies that used sensor- and pattern-based recognition technologies without reporting on the volatile biomarkers analyzed were not included in this review. Nevertheless, we believe that our report allows for a comprehensive review of exhaled VOC-based predictive models for distinguishing between benign PNs and lung cancer.

Conclusion

Exhaled 2-butanone, 3-hydroxy-2-butanone, and 2-hydroxyacetaldehyde might be significant breath markers for distinguishing between benign PNs and lung cancer. Moreover, predictive models that incorporate multiple factors alongside exhaled VOCs, such as demographic characteristics and radiological signs, have shown superior levels of performance compared with those based solely on exhaled VOCs. These observations highlight a conceivable role for VOCs as monitoring tools for PNs risk. However, the role of exhaled VOC-based predictive models in routine clinical practice has not yet been established, owing to various constraints associated with breath collection and analysis methods, as well as a general lack of external validation. Further investigations are therefore warranted to standardize the sample collection and measurement techniques for this approach, as well as to enhance the reliability and generalizability of such models.

Abbreviations

VOCs	Volatile Organic Compounds
PNs	Pulmonary Nodules
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
NOS	Newcastle-Ottawa Scale
PROBAST	Prediction Model Risk of Bias Assessment Tool
AUC	Area Under the Curve
NLST	National Lung Cancer Screening Trial
LDCT	Low-dose Computed Tomography
PROSPERO	The International Prospective Register of Systematic Reviews
CNKI	The Chinese National Knowledge Infrastructure
CBM	China Biology Medicine disc
MeSH	The Medical Subject Headings
GC-MS	Gas Chromatography-Mass Spectrometry
FT-ICR-MS	Fourier-Transform Ion Cyclotron Resonance Mass Spectrometry
HPPI-TOFMS	High-Pressure Photon Ionization Time-Of-Flight Mass Spectrometry
SPME	Solid-Phase Micro-Extraction
TD	Thermal Desorption
NIST Library	National Institute of Standards and Technology Library
LGOCV	Leave-Group-Out Cross-Validation
LOOCV	Leave-One-Out Cross-Validation
ERS	European Respiratory Society

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12890-024-03374-2.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

Author contributions

Weijuan Gong and Guangyu Lu had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.Concept and design: Weijuan Gong and Guangyu Lu. Acquisition, analysis, or interpretation of data: All authors. Drafting of the manuscript: Zhixia Su, Weijuan Gong and Guangyu Lu. Critical review of the manuscript for important intellectual content: All authors. Statistical analysis: Zhixia Su, Xiaoping Yu, Yuhang He, Taining Sha, Yujian Tao, Hong Guo, Liting Liao, Yanyan Zhang, and Guotao Lu. Administrative, technical, or material support: Weijuan Gong, Guangyu Lu, Xiaoping Yu, Yujian Tao, Hong Guo, Yanyan Zhang, and Guotao Lu. Supervision: Weijuan Gong and Guangyu Lu.

Funding

This study was supported by 2022 Jiangsu Provincial Science and Technology Programme Special Funds (Key R&D Programme for Social Development) [Grant no. BE2022775]; the National Natural Science Foundation of China [Grant no. 72374178]; the National Natural Science Foundation of China [Grant no. 71904165]; the Postgraduate Research and Practice Innovation Programme of Jiangsu Province [Grant no KYCX24_3852].

Data availability

All data generated or analyzed in the course of this study are included in these published articles and their supplementary information files.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Health Management Center, Affiliated Hospital of Yangzhou University, Yangzhou University, Yangzhou, Jiangsu 225012, China
[2]School of Public Health, Medical College of Yangzhou University, Yangzhou University, Yangzhou, Jiangsu 225009, China
[3]School of Nursing, Medical College of Yangzhou University, Yangzhou University, Yangzhou, Jiangsu 225009, China
[4]Department of Thoracic Surgery, Affiliated Hospital of Yangzhou University, Yangzhou, Jiangsu 225012, China
[5]Department of Respiratory and Critical Care Medicine, Affiliated Hospital of Yangzhou University, Jiangsu, Yangzhou 225012, China
[6]Department of Basic Medicine, Medical College of Yangzhou University, Yangzhou University, Yangzhou, Jiangsu 225009, China
[7]Testing Center of Yangzhou University, Yangzhou University, Yangzhou, Jiangsu 225009, China
[8]Yangzhou Key Laboratory of Pancreatic Disease, Institute of Digestive Diseases, Affiliated Hospital of Yangzhou University, Yangzhou University, Yangzhou, Jiangsu 225012, China
[9]Pancreatic Center, Department of Gastroenterology, Affiliated Hospital of Yangzhou University, Yangzhou University,, Yangzhou, Jiangsu 225012, China

Received: 18 April 2024 / Accepted: 29 October 2024
Published online: 02 November 2024

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and Mortality Worldwide for 36 cancers in 185 countries. Cancer J Clin. 2021;71(3):209–49. https://doi.org/10.3322/caac.21660.
2. Anderson MM, Hazen SL, Hsu FF, et al. Human neutrophils employ the myeloperoxidase-hydrogen peroxide-chloride system to convert hydroxy-amino acids into glycolaldehyde, 2-hydroxypropanal, and acrolein. A mechanism for the generation of highly reactive alpha-hydroxy and alpha,beta-unsaturated aldehydes by phagocytes at sites of inflammation. J Clin Invest. 1997;99(3):424–32. https://doi.org/10.1172/JCI119176.
3. Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends. Nat Rev Clin Oncol. 2023;20(9):624–39. https://doi.org/10.1038/s41571-023-00798-3.
4. Hirsch FR, Franklin WA, Gazdar AF, et al. Early detection of lung cancer: clinical perspectives of recent advances in biology and radiology. Clin Cancer Res. 2001;7(1):5–22. https://www.ncbi.nlm.nih.gov/pubmed/11205917. Published 2001/02/24.
5. Flehinger BJ, Melamed MR, Zaman MB, et al. Early lung cancer detection: results of the initial (prevalence) radiologic and cytologic screening in the Memorial Sloan-Kettering study. Am Rev Respir Dis. 1984;130(4):555–60. https://doi.org/10.1164/arrd.1984.130.4.555.
6. National Lung Screening Trial, Research T, Aberle DR, Adams AM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011;365(5):395–409. https://doi.org/10.1056/NEJMoa1102873.
7. Paci E, Puliti D, Lopes Pegna A, et al. Mortality, survival and incidence rates in the ITALUNG randomised lung cancer screening trial. Thorax. 2017;72(9):825–31. https://doi.org/10.1136/thoraxjnl-2016-209825.
8. Pastorino U, Silva M, Sestini S, et al. Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy. Ann Oncol. 2019;30(10):1672. https://doi.org/10.1093/annonc/mdz169.
9. Becker N, Motsch E, Trotter A, et al. Lung cancer mortality reduction by LDCT screening-results from the randomized German LUSI trial. Int J Cancer. 2020;146(6):1503–13. https://doi.org/10.1002/ijc.32486.
10. Gould MK, Tang T, Liu IL, et al. Recent trends in the identification of Incidental Pulmonary nodules. Am J Respir Crit Care Med. 2015;192(10):1208–14. https://doi.org/10.1164/rccm.201505-0990OC.
11. Harzheim D, Eberhardt R, Hoffmann H, et al. The Solitary Pulmonary Nodule. Respiration. 2015;90(2):160–72. https://doi.org/10.1159/000430996.
12. Kahn N, Kuner R, Eberhardt R, et al. Gene expression analysis of endobronchial epithelial lining fluid in the evaluation of indeterminate pulmonary nodules. J Thorac Cardiovasc Surg. 2009;138(2):474–9. https://doi.org/10.1016/j.jtcvs.2009.04.024.
13. Wahidi MM, Herth FJ, Ernst A. State of the art: interventional pulmonology. Chest. 2007;131(1):261–74. https://doi.org/10.1378/chest.06-0975.
14. Beigelman-Aubry C, Hill C, Grenier PA. Management of an incidentally discovered pulmonary nodule. Eur Radiol. 2007;17(2):449–66. https://doi.org/10.1007/s00330-006-0399-7.
15. Zhou M, Liu Y, Duan Y. Breath biomarkers in diagnosis of pulmonary diseases. Clin Chim Acta. 2012;413(21–22):1770–80. https://doi.org/10.1016/j.cca.2012.07.006.
16. de Lacy Costello B, Amann A, Al-Kateb H, et al. A review of the volatiles from the healthy human body. J Breath Res. 2014;8(1):014001. https://doi.org/10.1088/1752-7155/8/1/014001.
17. Wang P, Huang Q, Meng S, et al. Identification of lung cancer breath biomarkers based on perioperative breathomics testing: a prospective observational study. EClinicalMedicine. 2022;47:101384. https://doi.org/10.1016/j.eclinm.2022.101384.
18. Zou YC, Wang Y, Jiang ZL, et al. Breath profile as composite biomarkers for lung cancer diagnosis. Lung Cancer. 2021;154:206–13. https://doi.org/10.1016/j.lungcan.2021.01.020.
19. Tsou PH, Lin ZL, Pan YC, et al. Exploring volatile Organic compounds in Breath for High-Accuracy Prediction of Lung Cancer. Cancers (Basel). 2021;13(6). https://doi.org/10.3390/cancers13061431.
20. Peled N, Hakim M, Bunn PA Jr., et al. Non-invasive breath analysis of pulmonary nodules. J Thorac Oncology: Official Publication Int Association Study Lung Cancer. 2012;7(10):1528–33. https://doi.org/10.1097/JTO.0b013e3182637d5f.
21. Bousamra M 2nd, Schumer E, Li M, et al. Quantitative analysis of exhaled carbonyl compounds distinguishes benign from malignant pulmonary disease. J Thorac Cardiovasc Surg. 2014;148(3):1074–80. https://doi.org/10.1016/j.jtcvs.2014.06.006. discussion 1080 – 1071.
22. Fu XA, Li M, Knipp RJ, et al. Noninvasive detection of lung cancer using exhaled breath. Cancer Med. 2014;3(1):174–81. https://doi.org/10.1002/cam4.162.
23. Schumer EM, Trivedi JR, van Berkel V, et al. High sensitivity for lung cancer detection using analysis of exhaled carbonyl compounds. J Thorac Cardiovasc Surg. 2015;150(6):1517–22. https://doi.org/10.1016/j.jtcvs.2015.08.092. discussion 1522 – 1514.
24. Chen X, Muhammad KG, Madeeha C, et al. Calculated indices of volatile organic compounds (VOCs) in exhalation for lung cancer screening and early detection. Lung Cancer. 2021;154:197–205. https://doi.org/10.1016/j.lungcan.2021.02.006.
25. Poli D, Carbognani P, Corradi M, et al. Exhaled volatile organic compounds in patients with non-small cell lung cancer: cross sectional and nested short-term follow-up study. Respir Res. 2005;6(1):71. https://doi.org/10.1186/1465-9921-6-71.
26. Sakumura Y, Koyama Y, Tokutake H, et al. Diagnosis by Volatile Organic compounds in Exhaled Breath from Lung Cancer patients using support Vector Machine Algorithm. Sens (Basel). 2017;17(2). https://doi.org/10.3390/s17020287.
27. Phillips M, Bauer TL, Pass HI. A volatile biomarker in breath predicts lung cancer and pulmonary nodules. J Breath Res. 2019;13(3):036013. https://doi.org/10.1088/1752-7163/ab21aa.
28. Liao PQ. Research on lung cancer diagnosis model based on exhaled air analysis and machine learning [Thsis]. Sichuan: University of Electronic Science and Technology of China; 2021.
29. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. https://doi.org/10.1136/bmj.n71.
30. MacMahon H, Naidich DP, Goo JM, et al. Guidelines for management of Incidental Pulmonary nodules detected on CT images: from the Fleischner

Society 2017. Radiology. 2017;284(1):228–43. https://doi.org/10.1148/radiol.2017161659.
31. Modesti PA, Reboldi G, Cappuccio FP, et al. Panethnic differences in blood pressure in Europe: a systematic review and Meta-analysis. PLoS ONE. 2016;11(1):e0147601. https://doi.org/10.1371/journal.pone.0147601.
32. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: A Tool to assess risk of Bias and Applicability of Prediction Model studies: explanation and elaboration. Ann Intern Med. 2019;170(1):W1–33. https://doi.org/10.7326/M18-1377.
33. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: A Tool to assess the risk of Bias and Applicability of Prediction Model studies. Ann Intern Med. 2019;170(1):51–8. https://doi.org/10.7326/M18-1376.
34. Chen JH. Specific VOCs screening and application of multi-factor diagnostic model in lung cancer and pulmonary nodules [Thsis]. Shanghai: East China University of Science and Technology. 2019.
35. Chen YB. Preliminary study on the clinical value of exhaled volatile organic compounds in the diagnosis of lung cancer [Thsis]. Beijing: Chinese People's Liberation Army Medical College; 2017.
36. Hu YJ. Screening and diagnostic value of characteristic VOCs in lung cancer [Thsis]. Zhejaing: Zhejiang University; 2010.
37. Xu FJ. Study on characteristic gases in respiratory diagnosis of lung cancer [Thsis]. Zhejaing: Zhejiang University; 2008.
38. Orywal K, Szmitkowski M. Alcohol dehydrogenase and aldehyde dehydrogenase in malignant neoplasms. Clin Exp Med. 2017;17(2):131–9. https://doi.org/10.1007/s10238-016-0408-3.
39. Currie E, Schulze A, Zechner R, et al. Cellular fatty acid metabolism and cancer. Cell Metab. 2013;18(2):153–61. https://doi.org/10.1016/j.cmet.2013.05.017.
40. Cosnier F, Grossmann S, Nunge H, et al. Metabolism of inhaled methylethylketone in rats. Drug Chem Toxicol. 2018;41(1):42–50. https://doi.org/10.1080/01480545.2017.1289220.
41. Dietz FK, Rodriguez-Giaxola M, Traiger GJ, et al. Pharmacokinetics of 2-butanol and its metabolites in the rat. J Pharmacokinet Biopharm. 1981;9(5):553–76. https://doi.org/10.1007/bf01061026.
42. DiVincenzo GD, Kaplan CJ, Dedinas J. Characterization of the metabolites of methyl n-butyl ketone, methyl iso-butyl ketone, and methyl ethyl ketone in guinea pig serum and their clearance. Toxicol Appl Pharmacol. 1976;36(3):511–22. https://doi.org/10.1016/0041-008x(76)90230-1.
43. Otsuka M, Mine T, Ohuchi K, et al. A detoxication route for acetaldehyde: metabolism of diacetyl, acetoin, and 2,3-butanediol in liver homogenate and perfused liver of rats. J Biochem. 1996;119(2):246–51. https://doi.org/10.1093/oxfordjournals.jbchem.a021230.
44. Kovacic P, Somanathan R. Pulmonary toxicity and environmental contamination: radicals, electron transfer, and protection by antioxidants. Rev Environ Contam Toxicol. 2009;201:41–69. https://doi.org/10.1007/978-1-4419-0032-6_2.
45. Glorieux C, Liu S, Trachootham D, et al. Targeting ROS in cancer: rationale and strategies. Nat Rev Drug Discov. 2024;23(8):583–606. https://doi.org/10.1038/s41573-024-00979-4.
46. Al-Enezi KS, Alkhalaf M, Benov LT. Glycolaldehyde induces growth inhibition and oxidative stress in human breast cancer cells. Free Radic Biol Med. 2006;40(7):1144–51. https://doi.org/10.1016/j.freeradbiomed.2005.10.065.
47. Fuchs D, Jamnig H, Heininger P, et al. Decline of exhaled isoprene in lung cancer patients correlates with immune activation. J Breath Res. 2012;6(2):027101. https://doi.org/10.1088/1752-7155/6/2/027101.
48. Sukul P, Richter A, Junghanss C, et al. Origin of breath isoprene in humans is revealed via multi-omic investigations. Commun Biol. 2023;6(1):999. https://doi.org/10.1038/s42003-023-05384-y.
49. Xie SH, Dai W, Liu MX et al. Predictive value of volatile organic compounds in exhaled air for benign and malignant pulmonary nodules in < 50 years old. Chinese Clinical Journal of Thoracic and Cardiovascular Surgery. 2020;27(06):675–680. https://kns.cnki.net/kcms/detail/51.1492.R.20200512.1626.030.html
50. Malhotra J, Malvezzi M, Negri E, et al. Risk factors for lung cancer worldwide. Eur Respir J. 2016;48(3):889–902. https://doi.org/10.1183/13993003.00359-2016.
51. Silvestri GA, Goldman L, Burleson J, et al. Characteristics of persons screened for Lung Cancer in the United States: a Cohort Study. Ann Intern Med. 2022;175(11):1501–5. https://doi.org/10.7326/M22-1325.
52. Stoltzfus JC. Logistic regression: a brief primer. Acad Emerg Med. 2011;18(10):1099–104. https://doi.org/10.1111/j.1553-2712.2011.01185.x.
53. Bennett CM, Miller MB, Wolford GL. Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: an argument for multiple comparisons correction. NeuroImage. 2009;47:S125. https://doi.org/10.1016/S1053-8119(09)71202-9.
54. Ranstam J. Hypothesis-generating and confirmatory studies, Bonferroni correction, and pre-specification of trial endpoints. Acta Orthop. 2019;90(4):297. https://doi.org/10.1080/17453674.2019.1612624.
55. Chumbley JR, Friston KJ. False discovery rate revisited: FDR and topological inference using gaussian random fields. NeuroImage. 2009;44(1):62–70. https://doi.org/10.1016/j.neuroimage.2008.05.021.
56. LaValley MP. Logistic regression. Circulation. 2008;117(18):2395–9. https://doi.org/10.1161/circulationaha.106.682658.
57. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20(3):273–97. https://doi.org/10.1007/BF00994018.
58. Rai SN, Das S, Pan J, et al. Multigroup prediction in lung cancer patients and comparative controls using signature of volatile organic compounds in breath samples. PLoS ONE. 2022;17(11):e0277431. https://doi.org/10.1371/journal.pone.0277431.
59. Breiman L, Random, Forests. Mach Learn. 2001;45(1):5–32. https://doi.org/10.1023/A:1010933404324.
60. Ding X, Lin G, Wang P, et al. Diagnosis of primary lung cancer and benign pulmonary nodules: a comparison of the breath test and 18F-FDG PET-CT. Front Oncol. 2023;13:1204435. https://doi.org/10.3389/fonc.2023.1204435.
61. Li MX, Yang DK, Brock G, et al. Breath carbonyl compounds as biomarkers of lung cancer. Lung Cancer. 2015;90(1):92–7. https://doi.org/10.1016/j.lungcan.2015.07.005.
62. Callol-Sanchez L, Munoz-Lucas MA, Gomez-Martin O, et al. Observation of nonanoic acid and aldehydes in exhaled breath of patients with lung cancer. J Breath Res. 2017;11(2):026004. https://doi.org/10.1088/1752-7163/aa6485.
63. Phillips M, Altorki N, Austin JH, et al. Prediction of lung cancer using volatile biomarkers in breath. Cancer Biomark. 2007;3(2):95–109. https://doi.org/10.3233/cbm-2007-3204.
64. Phillips M, Altorki N, Austin JH, et al. Detection of lung cancer using weighted digital analysis of breath biomarkers. Clin Chim Acta. 2008;393(2):76–84. https://doi.org/10.1016/j.cca.2008.02.021.
65. Miekisch W, Kischkel S, Sawacki A, et al. Impact of sampling procedures on the results of breath analysis. J Breath Res. 2008;2(2):026007. https://doi.org/10.1088/1752-7155/2/2/026007.
66. Sukul P, Trefz P. Physio-metabolic monitoring via Breath employing real-time Mass Spectrometry: Importance, challenges, potentials, and Pitfalls. In: Weigl S, editor. Breath analysis: an Approach for Smart Diagnostics. Cham: Springer International Publishing; 2023. pp. 1–18.
67. Garcia-Marcos L, Edwards J, Kennington E, et al. Priorities for future research into asthma diagnostic tools: a PAN-EU consensus exercise from the European asthma research innovation partnership (EARIP). Clin Exp Allergy. 2018;48(2):104–20. https://doi.org/10.1111/cea.13080.
68. Schwoebel H, Schubert R, Sklorz M, et al. Phase-resolved real-time breath analysis during exercise by means of smart processing of PTR-MS data. Anal Bioanal Chem. 2011;401(7):2079–91. https://doi.org/10.1007/s00216-011-5173-2.
69. Haick H, Broza YY, Mochalski P, et al. Assessment, origin, and implementation of breath volatile cancer markers. Chem Soc Rev. 2014;43(5):1423–49. https://doi.org/10.1039/c3cs60329f.
70. van der Schee MP, Paff T, Brinkman P, et al. Breathomics in lung disease. Chest. 2015;147(1):224–31. https://doi.org/10.1378/chest.14-0781.
71. Mochalski P, King J, Unterkofler K, et al. Stability of selected volatile breath constituents in Tedlar, Kynar and Flexfilm sampling bags. Analyst. 2013;138(5):1405–18. https://doi.org/10.1039/c2an36193k.
72. Buszewski B, Kesy M, Ligor T, et al. Human exhaled air analytics: biomarkers of diseases. Biomed Chromatogr. 2007;21(6):553–66. https://doi.org/10.1002/bmc.835.
73. Agusti A, Bel E, Thomas M, et al. Treatable traits: toward precision medicine of chronic airway diseases. Eur Respir J. 2016;47(2):410–9. https://doi.org/10.1183/13993003.01359-2015.
74. McDonald VM, Fingleton J, Agusti A, et al. Treatable traits: a new paradigm for 21st century management of chronic airway diseases: treatable traits down under International Workshop report. Eur Respir J. 2019;53(5). https://doi.org/10.1183/13993003.02058-2018.
75. Wilson AD. Advances in electronic-nose technologies for the detection of volatile biomarker metabolites in the human breath. Metabolites. 2015;5(1):140–63. https://doi.org/10.3390/metabo5010140.
76. van der Sar IG, Wijbenga N, Nakshbandi G, et al. The smell of lung disease: a review of the current status of electronic nose technology. Respir Res. 2021;22(1):246. https://doi.org/10.1186/s12931-021-01835-4.

77. Baldini C, Billeci L, Sansone F et al. Electronic Nose as a Novel Method for Diagnosing Cancer: A Systematic Review. *Biosensors-Basel.* 2020;10(8). https://doi.org/ARTN 84.
78. Ramspek CL, Jager KJ, Dekker FW, et al. External validation of prognostic models: what, why, how, when and where? Clin Kidney J. 2021;14(1):49–58. https://doi.org/10.1093/ckj/sfaa188.

Publisher's note