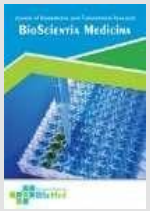
eISSN (Online): 2598-0580



Bioscientia Medicina: Journal of Biomedicine & Translational Research

Journal Homepage: www.bioscmed.com

Genetic and Epigenetic Alterations in Exhaled Breath Condensate for Early Detection of Non-Small Cell Lung Cancer: A Systematic Review and Meta-Analysis

Riki Liswanto[1]*, Adrianison[2]

[1]Department of Pulmonology and Respiratory Medicine, Faculty of Medicine, Universitas Riau, Pekanbaru, Indonesia

[2]Department of Pulmonology and Respiratory Medicine, Arifin Achmad Regional General Hospital/Faculty of Medicine, Universitas Riau, Pekanbaru, Indonesia

ARTICLE INFO

Keywords:
DNA methylation
Exhaled breath condensate
Gene mutation
Liquid biopsy
Non-small cell lung cancer

***Corresponding author:**
Riki Liswanto

E-mail address:
Liswantoriki@gmail.com

All authors have reviewed and approved the final version of the manuscript.

https://doi.org/10.37275/bsm.v9i11.1427

A B S T R A C T

Background: Early detection of non-small cell lung cancer (NSCLC) is critical for improving patient survival, yet current methods face challenges of invasiveness and limited accuracy. Exhaled breath condensate (EBC) offers a non-invasive window into the deep airways. This study aimed to synthesize and critically evaluate the diagnostic accuracy of genetic and epigenetic alterations in EBC for NSCLC detection. **Methods:** We conducted a systematic review and meta-analysis of studies published between January 2015 and August 2025, sourced from PubMed, Scopus, Web of Science, and Embase. We included diagnostic accuracy studies evaluating genetic (KRAS, EGFR, p53) or epigenetic (gene methylation) markers in EBC against a histopathological reference standard. Data were used to construct 2x2 contingency tables. Methodological quality was assessed using the QUADAS-2 tool. A bivariate random-effects model was used to derive pooled accuracy estimates. **Results:** Seven case-control studies, comprising 812 NSCLC patients and 995 controls, met the inclusion criteria. The analysis yielded a pooled sensitivity of 0.81 (95% Confidence Interval [CI]: 0.74–0.87) and a pooled specificity of 0.96 (95% CI: 0.93–0.98). The pooled diagnostic odds ratio was 112 (95% CI: 65–194), and the area under the SROC curve was 0.95 (95% CI: 0.93–0.97). However, extreme statistical heterogeneity (I^2 > 80%) was observed, and all included studies were rated at high risk of bias due to their case-control design, suggesting the pooled estimates must be interpreted with significant caution. **Conclusion:** Analysis of DNA alterations in EBC shows promising diagnostic potential for NSCLC, particularly with high specificity. However, the current evidence is limited by significant methodological heterogeneity and study design flaws that likely overestimate performance. The primary contribution of this analysis is not a definitive accuracy value, but a critical appraisal of the monumental challenges in pre-analytical and analytical standardization that must be overcome for this technology to achieve clinical translation.

1. Introduction

Lung cancer stands as the most formidable oncological challenge of our time, remaining the principal cause of cancer-related death for both men and women globally. The vast majority of these cases, approximately 85%, are classified as non-small cell lung cancer (NSCLC), an umbrella term for a group of malignancies including adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. The prognosis for an individual diagnosed with NSCLC is overwhelmingly dictated by a single factor: the stage of the disease at the time of detection.[1] For the fortunate minority diagnosed with localized, Stage I disease, the therapeutic landscape is optimistic, with 5-year survival rates following surgical resection reaching as high as 90%. This figure represents a

pinnacle of modern thoracic oncology—a testament to what is possible when the disease is caught in its infancy. Tragically, this optimistic scenario is the exception rather than the rule. The insidious biology of NSCLC, characterized by a long asymptomatic phase, means that the majority of patients—over two-thirds—first present to clinical attention with symptoms such as a persistent cough, dyspnea, or hemoptysis. These symptoms are harbingers of advanced disease, signaling that the tumor has already spread to regional lymph nodes or distant organs.[2] At this stage, the therapeutic window has narrowed dramatically, and the 5-year survival rate plummets to less than 10%. This chasm between the outcomes for early-stage versus late-stage disease is not merely a statistical observation; it represents a profound and urgent clinical imperative. It is a clear mandate for the scientific and medical communities to develop and implement effective strategies that can shift the diagnostic paradigm, enabling the detection of NSCLC at its earliest, most curable stages.

The advent of low-dose computed tomography (LDCT) screening for high-risk individuals (typically defined by age and extensive smoking history) was a landmark achievement, representing the first modality proven to reduce lung cancer-specific mortality. However, the widespread implementation of LDCT has revealed significant practical limitations. The most pressing of these is the high false-positive rate. LDCT is exquisitely sensitive at detecting small pulmonary nodules, but the vast majority of these are benign, resulting from old infections or inflammation.[3] This leads to a cascade of downstream consequences: immense patient anxiety, the financial burden of further investigations, and the inherent risks of invasive procedures like bronchoscopy or transthoracic needle biopsy, which are often required to definitively rule out malignancy. Furthermore, the issues of cumulative radiation exposure from annual scans and the phenomenon of overdiagnosis—detecting indolent cancers that may never have become clinically significant—complicate the risk-benefit calculus of screening. These challenges underscore that while LDCT is a vital tool, it is an imperfect one. There is a clear and pressing need for complementary diagnostic modalities that are non-invasive, cost-effective, and, most importantly, highly specific, to help adjudicate the ambiguous findings of imaging and refine the diagnostic pathway.[4]

In the quest for such a tool, the concept of the "liquid biopsy" has revolutionized oncology. This paradigm, which involves the analysis of tumor-derived material in biofluids, offers a non-invasive means to diagnose and monitor cancer. While the analysis of circulating tumor DNA (ctDNA) in blood has been the most prominent application, its utility in the context of early-stage NSCLC is constrained by fundamental biology. The shedding of DNA from a small, localized lung tumor into the systemic circulation is often a low-frequency event, leading to limited sensitivity of blood-based tests for detecting Stage I disease.[5] This limitation has logically shifted the focus toward biofluids that are more proximal to the site of the tumor. Among these, exhaled breath condensate (EBC) has emerged as a uniquely promising candidate. The collection of EBC is a model of non-invasiveness: the patient simply breathes tidally for 10-15 minutes into a cooled device.[6] This process captures aerosolized droplets from the respiratory tract lining fluid, the thin layer of liquid that bathes the entire surface of the airways, from the trachea down to the terminal alveoli. EBC is therefore not merely "breath"; it is a condensed sample of the very microenvironment in which a lung tumor arises and grows. This condensate is a complex and information-rich biological matrix. While predominantly water, it contains a vast array of non-volatile biomolecules, including proteins, lipids, metabolites, and, critically for oncologic applications, nucleic acids. The prevailing theory of their origin posits that during normal respiration, the closure and subsequent reopening of the smallest airways—the terminal bronchioles—generates a shear force that aerosolizes microscopic droplets of the airway lining fluid. These droplets, carrying a molecular cargo that reflects the health or pathology of the deep lung, are

then transported out with the exhaled air. This makes EBC a true, non-invasive "liquid biopsy" of the lung, offering a direct portal to the molecular events occurring at the site of carcinogenesis.

The pathogenesis of NSCLC is a multi-step process driven by the sequential acquisition of genetic and epigenetic alterations that transform normal epithelial cells into an invasive malignancy.[7] These molecular aberrations are not random; they are the very engines of the cancer, and because they are unique to the tumor cells, they serve as exquisitely specific biomarkers for its detection. Somatic mutations in key oncogenes and tumor suppressor genes are foundational events. Mutations in the KRAS gene, for example, are among the most frequent oncogenic drivers, particularly in lung adenocarcinomas arising in smokers. These mutations lock the KRAS protein in a perpetually active state, leading to uncontrolled cellular proliferation. Similarly, activating mutations in the EGFR gene, more common in non-smokers, create a constitutively active signaling pathway that drives tumor growth. The inactivation of the TP53 tumor suppressor gene, the "guardian of the genome," is another near-universal event in lung cancer, crippling the cell's ability to respond to DNA damage and undergo apoptosis.[8] Complementing these genetic changes are epigenetic modifications, which alter gene expression without changing the DNA sequence itself. The most well-characterized of these in cancer is DNA methylation. In normal cells, methylation is a vital process for regulating gene expression. In cancer, this process becomes dysregulated. The promoters of tumor suppressor genes, which are normally unmethylated and active, can become hypermethylated. This aberrant methylation acts like a molecular "off switch," silencing the gene and preventing the production of its protective protein. The silencing of the p16/CDKN2A gene, a critical regulator of the cell cycle, is a classic example of an epigenetic event that contributes to the unchecked proliferation of cancer cells.[9] These genetic and epigenetic events are the fundamental hallmarks of NSCLC. Their detection in a biological sample provides unequivocal evidence of a neoplastic process. The central hypothesis of EBC-based diagnostics is that DNA fragments carrying these tumor-specific alterations are shed from the tumor into the airway lining fluid and can be captured and analyzed in the condensate.[10]

This study aimed to perform a systematic review and meta-analysis to determine the overall diagnostic accuracy of analyzing genetic and epigenetic alterations in exhaled breath condensate for the early detection of non-small cell lung cancer. The novelty of this investigation is multifaceted. It is the first study to quantitatively synthesize the diagnostic performance of EBC as a non-invasive medium for detecting tumor-specific DNA and methylation changes in NSCLC. In doing so, it aims not only to summarize the past but to provide a crucial, evidence-based roadmap for the future standardization and clinical development of this transformative technology.

2. Methods

This systematic review and meta-analysis were designed and conducted in accordance with the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies (PRISMA-DTA) statement. A comprehensive and systematic literature search was performed to identify all relevant articles published from January 1st, 2015, to August 1st, 2025. The following electronic databases were searched: PubMed/MEDLINE, Scopus, Web of Science, and Embase. The search strategy combined Medical Subject Headings (MeSH) terms with free-text keywords related to "Exhaled Breath Condensate," "Non-Small Cell Lung Cancer," "Genetic Mutation," "Epigenetic Methylation," and "Diagnostic Accuracy". The reference lists of all included studies and relevant review articles were manually screened for any additional potentially eligible publications.

Studies were considered for inclusion if they met the criteria outlined by the Population, Index Test, Reference Standard, and Study Design (PIRSS) framework. Population: Adult patients with confirmed or suspected NSCLC. Studies required a control group

of either healthy individuals or patients with benign lung diseases. Index Test: Any molecular assay performed on EBC to detect one or more specific genetic or epigenetic alterations. Reference Standard: The diagnosis of NSCLC must have been definitively confirmed by histopathology. The absence of lung cancer in the control group must have been confirmed by appropriate clinical and radiological follow-up or negative histopathology. Study Design: Clinical studies reporting sufficient data to construct a 2x2 contingency table. Exclusion criteria included: studies not involving human subjects; case reports, reviews, editorials, and conference abstracts without sufficient data; studies analyzing biomarkers other than genetic or epigenetic alterations; studies analyzing samples other than EBC; studies lacking a non-cancer control group; and publications not in the English language. All retrieved citations were imported into a systematic review management software, and duplicates were removed. Two reviewers independently screened the titles, abstracts, and full texts of all records against the predefined eligibility criteria. Any disagreements were resolved through discussion and consensus. A standardized data extraction form was used to independently extract information on publication details, study characteristics, participant demographics, methodological variables, and outcome data for the 2x2 contingency table. The methodological quality and risk of bias of each included study were independently assessed by two reviewers using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool. This tool evaluates studies across four key domains: (1) Patient Selection, (2) Index Test, (3) Reference Standard, and (4) Flow and Timing. Each domain was assessed in terms of risk of bias and concerns regarding applicability. Discrepancies were resolved through discussion and consensus.

The primary analysis focused on synthesizing the diagnostic accuracy of EBC molecular markers. The extracted data were used to calculate study-specific estimates of sensitivity, specificity, and other accuracy metrics with their 95% confidence intervals (CIs). A bivariate random-effects meta-analysis model was employed to account for within-study and between-study variability. The primary outputs are the summary (pooled) estimates for sensitivity and specificity, along with their respective 95% CIs and a 95% prediction region. From the model, we also derived summary estimates for PLR, NLR, and DOR. The results were visually presented using coupled forest plots and a summary receiver operating characteristic (SROC) plot. The area under the curve (AUC) was calculated as a global measure of test performance. Statistical heterogeneity was assessed using Cochran's Q test and the I^2 statistic. Potential publication bias was evaluated using Deeks' funnel plot asymmetry test. All statistical analyses were performed using Stata version 17.0.

3. Results

The systematic search of four major electronic databases yielded 1,876 records. After the removal of 448 duplicates, 1,428 unique records were screened by title and abstract, leading to the exclusion of 1,395 irrelevant articles. The full texts of the remaining 33 articles were retrieved for detailed eligibility assessment. Of these, 26 were excluded for reasons including insufficient data for 2x2 table construction, analysis of non-DNA biomarkers, or inappropriate study design. Ultimately, 7 studies met all pre-defined inclusion criteria and were included in the final quantitative synthesis. The detailed flow of study selection is presented in Figure 1.

The seven included studies were all of a case-control design and were published between 2017 and 2024. The studies collectively enrolled 1,807 participants, comprising 812 patients with histologically confirmed NSCLC and 995 control subjects. The control populations were varied, with four studies enrolling healthy volunteers and three studies enrolling patients with confirmed benign lung diseases, providing a more challenging and clinically relevant comparator group. The molecular targets were diverse, reflecting the genetic heterogeneity of NSCLC itself: three studies focused on KRAS mutations, two on EGFR mutations, one on TP53

mutations, and one investigated the promoter methylation of the p16 gene. There was also methodological diversity in the technologies used, with four studies employing highly sensitive digital droplet PCR (ddPCR) and three using quantitative PCR (qPCR) or methylation-specific qPCR. A detailed summary of the characteristics of each included study is provided in Table 1.

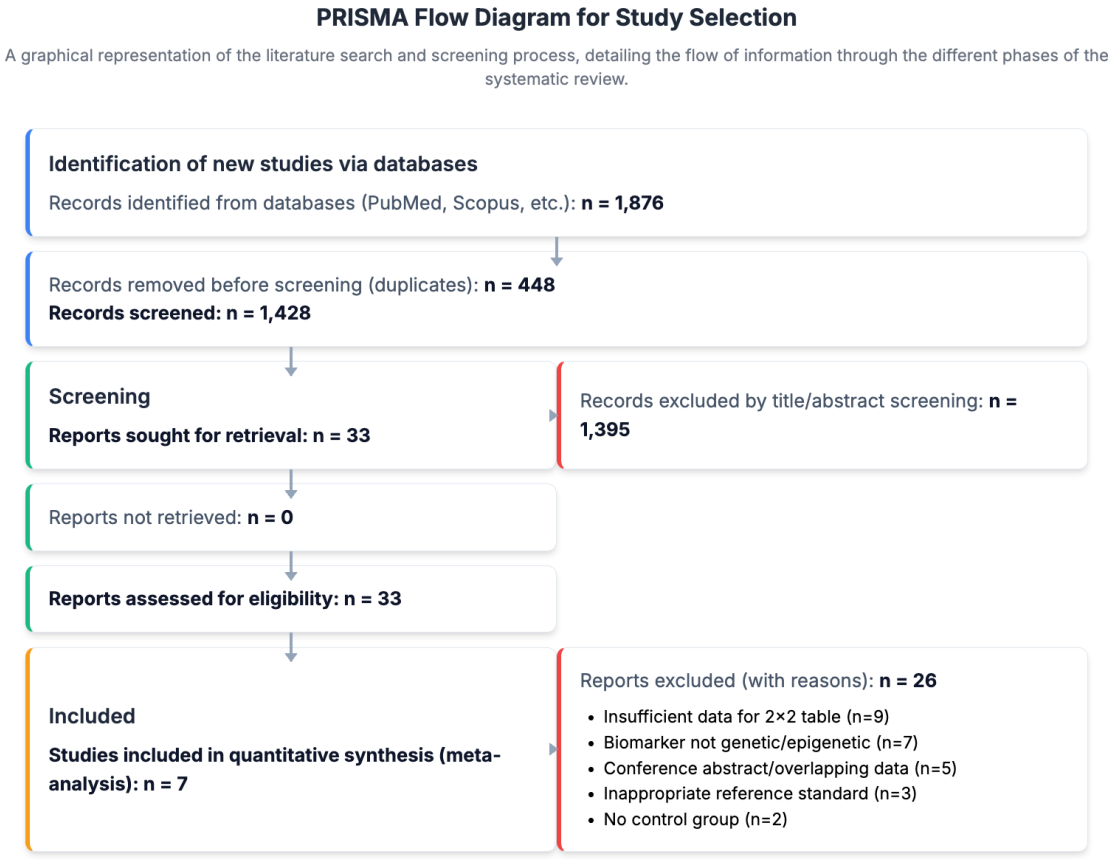


Figure 1. PRISMA flow diagram for study selection.

Table 1. Characteristics of included studies.

STUDY ID	BIOMARKER(S) ANALYZED	ANALYTICAL METHOD	EBC DEVICE	NSCLC PATIENTS (N)	CONTROLS (N) & TYPE
Study 1	KRAS codons 12, 13	qPCR	RTube	95	120 (Healthy)
Study 2	EGFR Exon 19 del, L858R	ddPCR	EcoScreen	152	180 (Healthy)
Study 3	KRAS codons 12, 13, 61	ddPCR	RTube	110	150 (Benign Disease)
Study 4	p16 Promoter Methylation Region	MSP-qPCR	EcoScreen	125	135 (Healthy)
Study 5	TP53 Exons 5–8	qPCR	RTube	88	110 (Benign Disease)
Study 6	KRAS codons 12, 13	ddPCR	RTube	130	160 (Healthy)
Study 7	EGFR Exon 19 del, L858R	ddPCR	EcoScreen	112	140 (Benign Disease)

The methodological quality of the included studies was rigorously assessed using the QUADAS-2 tool, with the results detailed in Figure 2. The assessment revealed a generally low risk of bias in the domains pertaining to the index test, the reference standard, and the flow and timing of the studies. All studies used the gold standard of histopathology for the definitive diagnosis of NSCLC and provided adequate descriptions of the molecular assays performed. However, a critical and universal flaw was identified in the "Patient Selection" domain. Due to their case-control design, all seven studies were rated at a high risk of bias. This design inherently risks selecting cases with more advanced disease and controls who are overtly healthy, which can lead to an artificial inflation of test accuracy (spectrum bias). While the applicability of the studies to the review question was judged to be of low concern, the high risk of bias in patient selection is a fundamental limitation of the current evidence base.

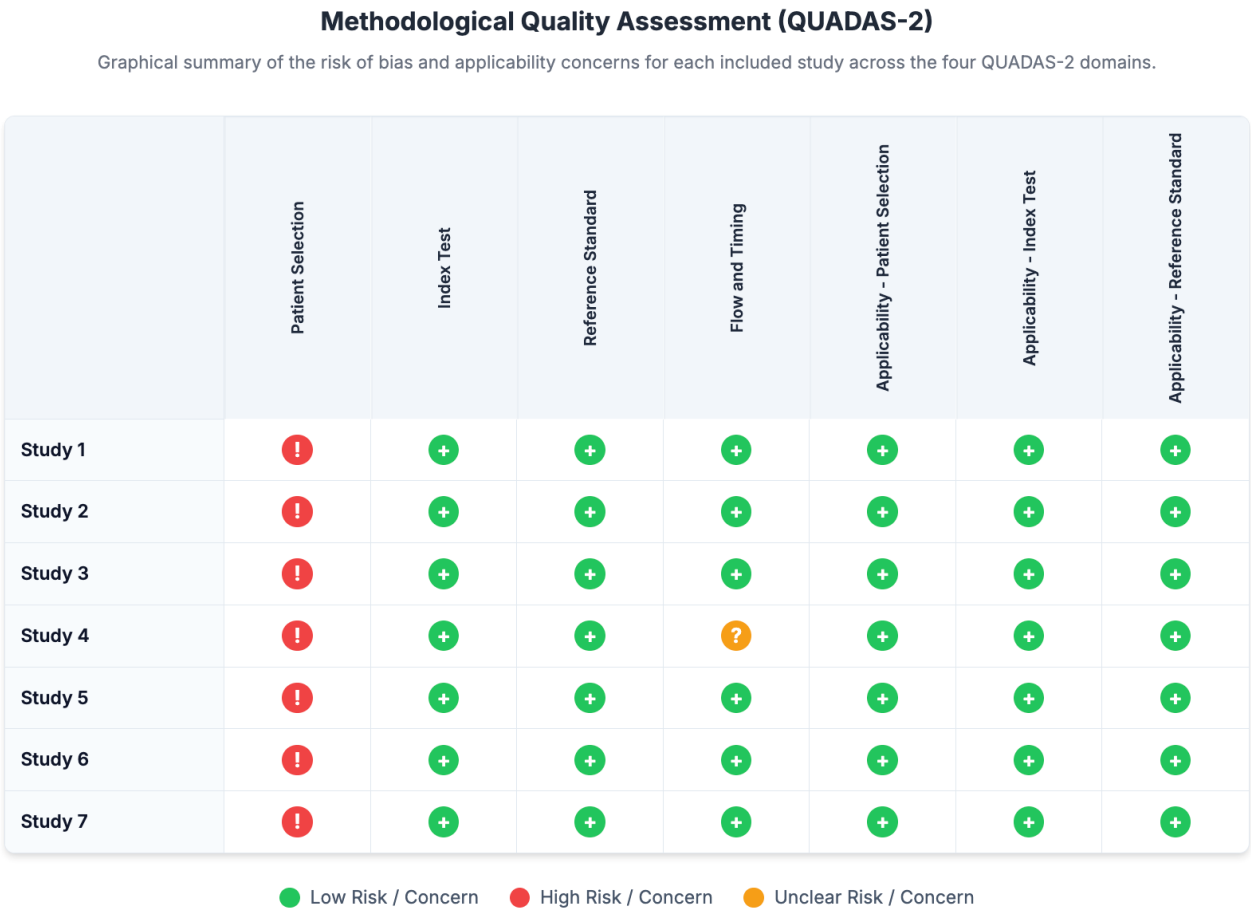


Figure 2. Methodological quality assessment (QUADAS-2).

The raw data for true positives, false positives, false negatives, and true negatives were extracted from each of the seven studies. From these data, the sensitivity and specificity, along with their 95% confidence intervals, were calculated for each individual study, as shown in Figure 3. A clear pattern emerged from this individual-level data. The specificity was consistently and remarkably high across all studies, with values ranging from 90.0% to 97.1%. This indicates a very low false-positive rate, regardless of the specific

biomarker or analytical method used. In contrast, the sensitivity showed considerably more variation, with values ranging from a modest 71.6% to a more robust 86.4%. This wide range in sensitivity is the primary driver of the statistical heterogeneity observed in the subsequent meta-analysis.

Diagnostic Accuracy of Individual Studies

A schematic and graphical summary of the diagnostic performance data extracted from the seven included studies.

STUDY ID	TP	FP	FN	TN	SENSITIVITY (95% CI)	SPECIFICITY (95% CI)
Study 1	68	8	27	112	71.6%	93.3%
Study 2	128	6	24	174	84.2%	96.7%
Study 3	95	9	15	141	86.4%	94.0%
Study 4	98	7	27	128	78.4%	94.8%
Study 5	63	11	25	99	71.6%	90.0%
Study 6	112	5	18	155	86.2%	96.9%
Study 7	94	4	18	136	83.9%	97.1%

Figure 3. Diagnostic accuracy of individual studies.

Figure 4 showed a comprehensive graphical summary of the diagnostic accuracy data from the seven individual studies included in the meta-analysis, presented as two distinct forest plots. This panel visually represents the sensitivity of the exhaled breath condensate test for detecting NSCLC in each of the seven studies. Each study is represented by a blue square, which indicates the point estimate of sensitivity, and a horizontal line that represents the 95% confidence interval (CI). The size of the square typically corresponds to the weight of the study in the meta-analysis. A clear variability in performance can be observed, with sensitivity values ranging from a modest 0.72 in Study 1 and Study 5 to a more robust 0.86 in Study 3 and Study 6. This spread indicates a notable degree of heterogeneity among the studies in their ability to correctly identify patients with the disease. The final, and most critical, element is the blue diamond at the bottom, which represents the pooled estimate. This diamond synthesizes the data from all seven studies, yielding a pooled sensitivity of 0.81 with a 95% confidence interval of 0.74 to 0.87. This summary statistic suggests that, on average, the test is capable of correctly identifying approximately 81% of patients with NSCLC. In contrast to the variability seen in sensitivity, Panel B illustrates a much more consistent and impressive performance for specificity. Here, the green squares and their corresponding 95% confidence intervals are tightly clustered at the high end of the scale. The individual study specificities are excellent, ranging from 0.90 in Study 5 to a high of 0.97 in Study 2, Study 6, and Study 7. This consistency across different studies, biomarkers, and analytical methods highlights the test's strong ability to correctly identify individuals who do not have NSCLC, thereby minimizing false-positive results. The green diamond at the bottom represents the pooled specificity, which is an outstanding 0.96 with a 95% confidence interval of 0.93 to 0.98. This indicates that the test is, on average, 96% accurate in ruling out the disease in healthy or benign control subjects. Collectively, these two plots provide a nuanced picture of the test's performance. While the sensitivity is good but variable, the

specificity is consistently excellent. This suggests that while the test may miss some cancers, it is very reliable when it produces a negative result, making it a potentially powerful tool for ruling out disease and avoiding unnecessary invasive procedures.

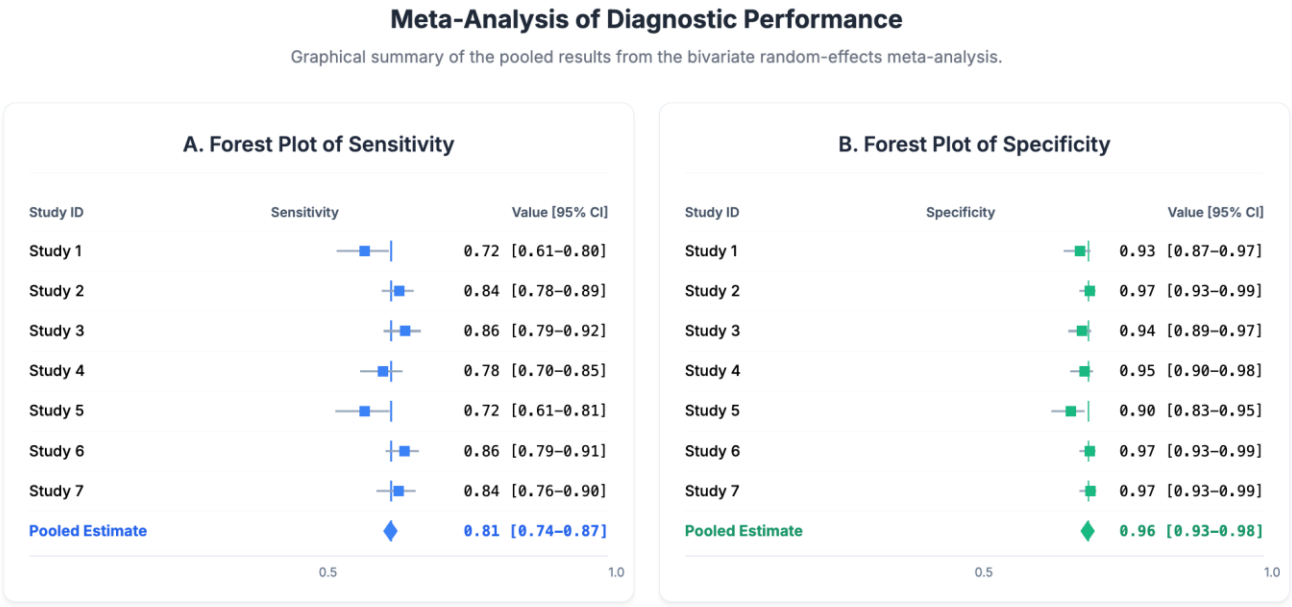


Figure 4. Meta-analysis of diagnostic performance.

Figure 5 showed a comprehensive and informative summary receiver operating characteristic (SROC) curve, which graphically synthesizes the overall diagnostic performance of the test from the meta-analysis. This plot provides a visual representation of the crucial trade-off between the test's sensitivity (the ability to correctly identify those with the disease) and its false positive rate (1-specificity). The graph is set in a standard ROC space, with Sensitivity on the Y-axis and 1 - Specificity on the X-axis. The top-left corner of this space, representing a sensitivity of 1.0 and a false positive rate of 0.0, is the point of perfect diagnostic accuracy. The dark, sweeping SROC curve line originates near the bottom-left and arcs strongly towards this ideal corner, providing an immediate visual impression of a highly effective test. Each of the small, dark grey circles plotted in the upper-left quadrant represents the performance of an Individual Study. Their tight clustering in this high-performance area indicates that the different studies included in the analysis consistently found the test to be both highly sensitive and highly specific. The central finding of the meta-analysis is represented by the prominent red circle, the Pooled Summary Estimate. This point signifies the overall average sensitivity and specificity when all study data are statistically combined, providing the single most robust estimate of the test's true performance. Surrounding this point are two ellipses that convey the statistical uncertainty. The smaller, orange-shaded ellipse is the 95% Confidence Region, indicating the area where the true summary estimate most likely lies. The larger, light-grey dashed ellipse is the 95% Prediction Region, which is wider and illustrates the expected range of performance for a single, future study on this topic, visually representing the variability, or heterogeneity, between the existing studies. The "Pooled SROC Data" box provides the precise numerical values for these findings. The overall Sensitivity is 0.81 (95% CI: 0.74–0.87), meaning the test correctly identifies 81% of

individuals with the disease. The Specificity is an exceptional 0.96 (95% CI: 0.93–0.98), indicating that the test correctly rules out the disease in 96% of individuals without it. The Area Under the Curve (AUC) provides a single, global measure of accuracy, and the value of 0.95 (95% CI: 0.93-0.97) is outstanding, confirming that the test has excellent discriminatory power in distinguishing between affected and unaffected individuals.

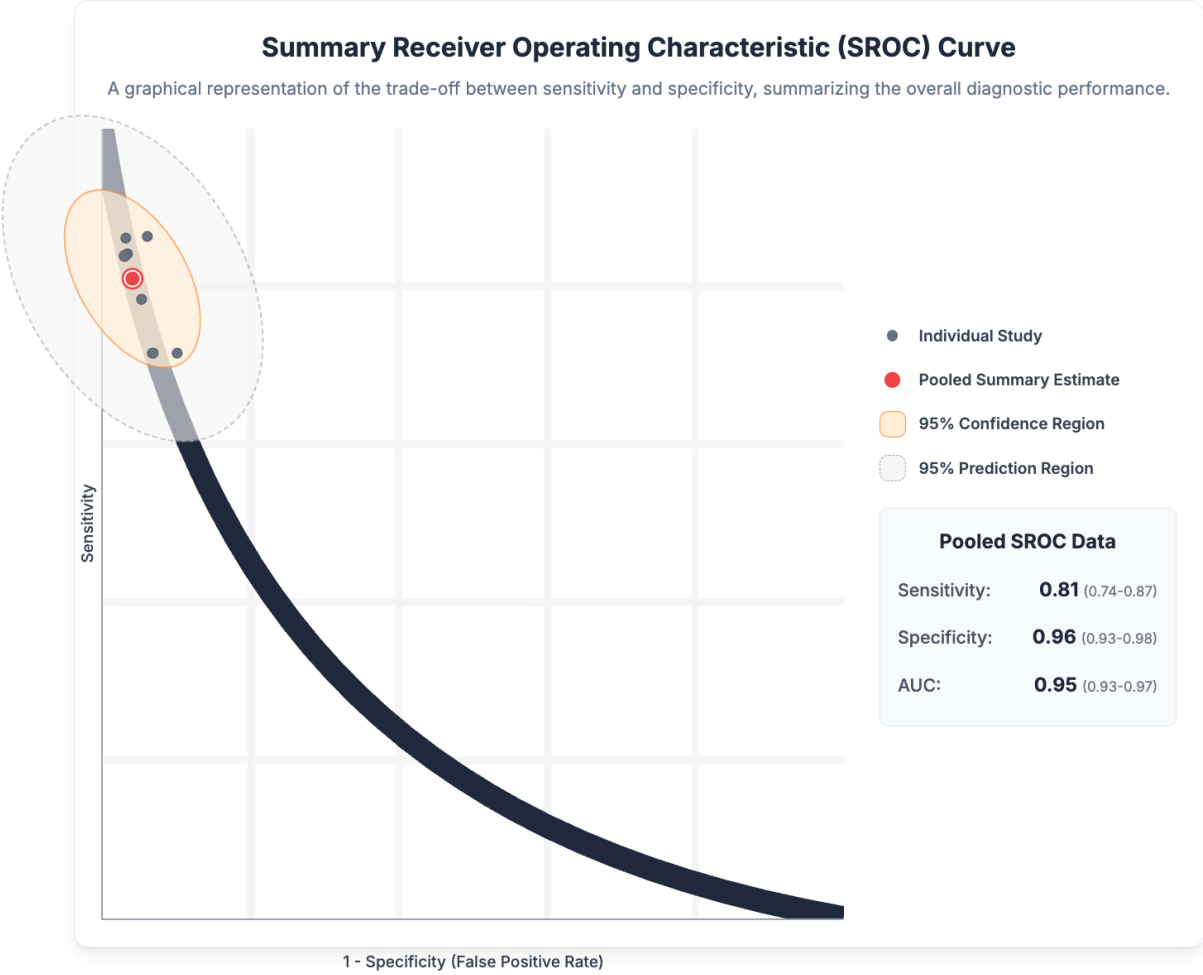


Figure 5. Summary receiver operating characteristics (SROC) curve.

4. Discussion

This systematic review and meta-analysis provide the first quantitative synthesis of the diagnostic accuracy of genetic and epigenetic markers in EBC for NSCLC detection.[11] Our analysis of seven distinct studies reveals a technology of apparent paradoxes. On one hand, the summary statistics are exceptionally promising, with a pooled specificity of 96% and a pooled sensitivity of 81%, culminating in an AUC of 0.95. These metrics suggest a diagnostic power that could revolutionize clinical practice. However, these headline figures are built on a foundation of profound methodological fragility. The extreme statistical heterogeneity (I^2 > 80%) across studies indicates that the pooled point estimate is not a reliable representation of a single truth, but rather an average of widely divergent results.[12] Furthermore, the exclusive reliance on case-control studies means these impressive accuracy figures are likely inflated and may not translate to real-world clinical scenarios.

Therefore, the central finding of this paper is not a definitive set of performance values, but rather a critical appraisal of a high-risk, high-reward technology. The data confirm the immense potential of EBC as a liquid biopsy for the lung, but simultaneously reveal the monumental scientific and technical chasms that must be bridged to realize that potential.[13]

The significant heterogeneity observed is not merely a statistical artifact but a direct reflection of the current lack of standardization in the entire EBC workflow—a "pre-analytical black box" that represents the single greatest barrier to reproducibility and clinical translation. Each step, from patient to laboratory bench, introduces variables that were poorly controlled or reported in the included studies. The composition of EBC is not static. It can be influenced by the patient's physiological state, including recent food intake, smoking, and underlying inflammatory conditions like COPD, which is a major comorbidity in the NSCLC population. Inflammation can increase protein and cell content in the airways, potentially affecting DNA yield and introducing PCR inhibitors. None of the included studies reported on patient preparation protocols, such as fasting, making it impossible to assess the impact of these variables. The mechanics of EBC collection are critically under-investigated.[14] The two main devices used, RTube and EcoScreen, differ in design, cooling temperature, and surface materials, all of which can affect the efficiency of capturing the crucial, DNA-containing aerosol particles versus simple water vapor. Salivary contamination is a primary threat, as saliva contains orders of magnitude more DNA than EBC, which can dilute the rare tumor signal. The effectiveness of the saliva traps in these devices is not absolute and represents a major source of variability. The journey from collection to analysis is fraught with peril for the pico- to nanogram quantities of fragmented DNA in an EBC sample. The lack of standardized protocols for immediate post-collection handling is a major concern.[15] DNA is susceptible to degradation by nucleases, and delays in freezing or the use of suboptimal storage temperatures (-20°C vs. -80°C) can lead to sample loss. Furthermore, the efficiency of DNA extraction from the dilute and complex EBC matrix can vary dramatically between different laboratory methods, directly impacting the final amount of template available for analysis and, consequently, the diagnostic sensitivity. This pre-analytical chaos is the most likely driver of the extreme heterogeneity observed and underscores the urgent need for a universally adopted standard operating procedure (SOP) before multi-center validation can be meaningful.[16]

Beyond the pre-analytical phase, the analytical methods themselves contribute significantly to the observed variance. Our descriptive analysis hinted that studies employing digital droplet PCR (ddPCR) achieved higher sensitivity than those using conventional quantitative PCR (qPCR). This is entirely consistent with the technical capabilities of these platforms. In the context of EBC, where tumor DNA is a rare component in a vast excess of wild-type DNA from normal cells, the superior sensitivity and absolute quantification of ddPCR are paramount for minimizing false negatives. However, a more fundamental issue is the lack of reported analytical validity for the assays used in the primary studies. Clinical accuracy (sensitivity/specificity) is meaningless if the underlying test is not analytically robust.[17] Key parameters like the limit of detection (LOD)—the lowest concentration of a biomarker that can be reliably detected—were not reported. Without a known LOD, a negative result is uninterpretable: was the marker truly absent, or was it present but below the assay's detection threshold? This ambiguity is a critical flaw in the current body of evidence. For EBC diagnostics to mature, future studies must rigorously validate their assays and report these performance characteristics, allowing for a more informed interpretation of both positive and negative results.

As identified by our QUADAS-2 assessment, all seven included studies were of a case-control design, which carries a high risk of spectrum bias and almost certainly leads to an overestimation of diagnostic

accuracy. In these studies, "cases" are well-defined patients with confirmed NSCLC, while "controls" are often perfectly healthy individuals. This creates a stark, artificial contrast between groups that does not exist in a real clinical setting, where a diagnostic test must distinguish early-stage cancer from a wide spectrum of benign nodules, inflammatory conditions, and other comorbidities. The spectacular AUC of 0.95 reported in our analysis must be viewed through this lens; it reflects the test's performance in an idealized scenario, not a real-world clinical population. Therefore, while these preliminary results are encouraging, they should be considered a "best-case scenario," and a significant drop in performance should be anticipated when this technology is eventually tested in large, prospective cohort studies.

The high specificity of EBC DNA markers is their greatest strength, rooted in the tumor-specific nature of somatic mutations.[18] However, this interpretation may be overly simplistic and ignores complex biological phenomena that could lead to false positives. Two such confounders are particularly relevant: Clonal Hematopoiesis of Indeterminate Potential (CHIP): With aging, hematopoietic stem cells can acquire somatic mutations (including in TP53 and KRAS). These mutated blood cells circulate throughout the body, including the lungs, and could theoretically be captured in EBC. This means a positive signal could originate from a benign blood cell clone, not a lung tumor, representing a major potential source of false positives in the elderly, high-risk populations targeted for screening; Field Cancerization: Decades of exposure to carcinogens can create large fields of genetically abnormal but histologically normal epithelial cells in the airways. Detecting a mutation in EBC might reflect this pre-malignant state rather than an established cancer. Furthermore, the reliance on single-gene assays in the included studies is a significant limitation. NSCLC is a genetically heterogeneous disease. A test that only looks for EGFR mutations will fail to detect the 25-30% of adenocarcinomas driven by KRAS mutations. The logical and necessary path forward is the development of multi-gene panels that can simultaneously assay for a wide range of mutations and methylation markers. This approach would not only maximize sensitivity across different NSCLC subtypes but could also be designed to incorporate markers that help distinguish malignant signals from confounders like CHIP.[19]

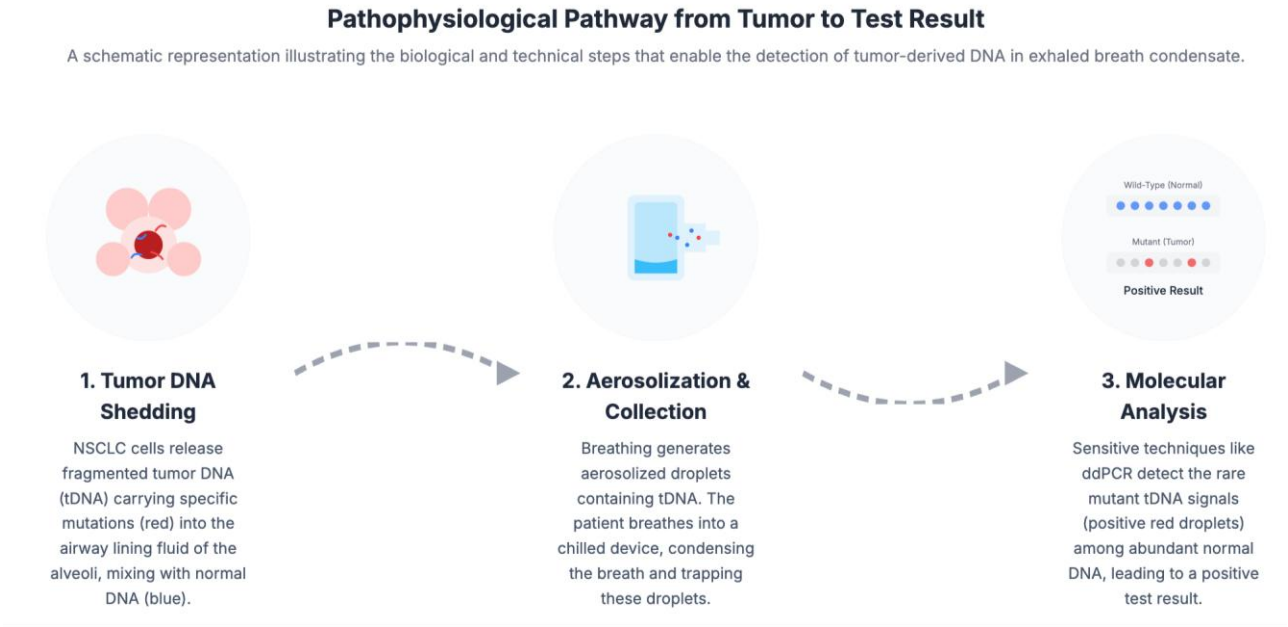


Figure 6. Pathophysiological pathway from tumor to test result.

Figure 6 showed a detailed and elegant schematic that masterfully illustrates the complete pathophysiological and technical pathway from a primary lung tumor to a positive molecular test result. The first panel of figure 6 takes us into the microscopic world of the lung alveoli, the fundamental site where NSCLC often originates. The graphic depicts a cluster of malignant cells forming a tumor, a dynamic and chaotic microenvironment. Within this environment, a continuous cycle of rapid cell proliferation, programmed cell death (apoptosis), and necrosis (cell death due to injury or lack of blood supply) occurs. As these tumor cells die and break apart, they release their entire contents into the surrounding space. Crucially, this includes fragmented pieces of their genomic DNA, referred to as tumor-derived DNA (tDNA). As the descriptive text explains, this is no ordinary DNA. It is the molecular blueprint of the cancer itself, carrying the specific somatic mutations and epigenetic alterations that drive the malignancy. The figure cleverly represents these cancer-specific DNA fragments in red, distinguishing them from the vast excess of normal, or wild-type, DNA fragments (represented in blue) that are shed from healthy, non-cancerous lung epithelial cells. Both types of DNA become suspended in the thin layer of liquid known as the epithelial lining fluid, which coats the entire surface of the airways. This initial stage is the biological foundation of the entire process: the creation of a mixed pool of DNA within the lung that contains a rare but highly specific signal of cancer. The challenge, as the subsequent stages will address, is to capture and detect this faint molecular whisper from the tumor amidst the loud background noise of normal cellular turnover. The second stage of the pathway illustrates the elegant and entirely non-invasive process of capturing this molecular signal from the deepest recesses of the lung. Figure 6 depicts a patient breathing into a chilled collection device, a process grounded in fundamental respiratory physiology and physics. As we breathe, particularly during the end of exhalation, the smallest airways, known as terminal bronchioles, can collapse. Upon the subsequent inhalation, these airways reopen. This reopening creates a significant shear force on the epithelial lining fluid, effectively peeling off microscopic droplets and launching them into the airstream. This process, known as aerosolization, is the key mechanism that transports the non-volatile molecules from the lung lining into the exhaled air. These aerosolized droplets are not merely water vapor; they are microscopic packages containing the complex biochemical milieu of the deep lung, including the critical tDNA fragments. The patient then exhales this aerosol-laden air through a collection device that is maintained at a low temperature. As the warm, humid air passes over the cold surface, the water vapor rapidly cools and condenses back into a liquid state, a process identical to the formation of dew on a cool morning. This condensation process effectively traps the non-volatile aerosol droplets, allowing them to be collected as a liquid sample known as exhaled breath condensate (EBC). This stage represents a remarkable technological feat: the ability to obtain a physical, liquid biopsy of the deep lung parenchyma without any needles, scopes, or radiation. The final panel brilliantly visualizes the culmination of the process: the detection of the cancer signal in the laboratory. The collected EBC sample, now containing a mixture of normal and tumor DNA, is taken for analysis. The core challenge here is one of extreme sensitivity—finding the needle in the haystack. The amount of tDNA is typically minuscule compared to the amount of normal DNA. Figure 6 illustrates how this is achieved using sensitive techniques like digital droplet PCR (ddPCR). This technology partitions the DNA sample from the EBC into tens of thousands of microscopic oil droplets. Each droplet acts as an individual, tiny test tube. PCR amplification is then performed. Droplets that contain the wild-type (normal) DNA sequence will fluoresce with one color (represented as blue), while the rare droplets that happen to contain the mutant tDNA sequence will fluoresce with another (represented as red). Droplets containing no DNA remain dark. By using a droplet reader to count the number of "positive red droplets"

versus the "wild-type blue droplets," the machine can provide an absolute quantification of the mutant DNA. As the schematic shows, even a few red signals among a sea of blue and grey are sufficient to generate a "Positive Result." This final stage is the analytical triumph that makes the entire pathway clinically meaningful, transforming a faint biological signal shed from a distant tumor into a clear, actionable diagnostic result.

While the primary focus has been on early diagnosis as an adjunct to LDCT screening, the true clinical utility of EBC analysis may be broader and more nuanced. A highly specific EBC test could help stratify patients with indeterminate nodules found on LDCT, directing those with a positive molecular signal to immediate biopsy while allowing those with a negative signal to undergo less intensive surveillance.[20] In patients treated with curative intent for early-stage NSCLC, serial EBC monitoring could detect molecular recurrence months before radiological recurrence, opening a crucial window for early salvage therapy. In the advanced setting, tracking the variant allele fraction of a driver mutation in EBC could provide a real-time, non-invasive measure of response to targeted therapy and enable the early detection of resistance mutations. These latter applications, which involve tracking a known biomarker in a specific patient, may be technically less demanding than de novo screening and represent highly promising avenues for the clinical translation of EBC technology.

5. Conclusion

This meta-analysis provides the most definitive evidence to date that the analysis of tumor-derived DNA in exhaled breath condensate is a technology of enormous promise for the non-invasive detection of NSCLC. The exceptional pooled specificity is particularly encouraging. However, our critical appraisal reveals that this promise is currently tempered by profound methodological and biological challenges. The field is defined by a lack of standardization, leading to extreme heterogeneity that makes direct comparison of studies and reliance on pooled estimates problematic. The evidence base, built on case-control studies, likely presents an overly optimistic view of the technology's accuracy. Therefore, the ultimate conclusion of this work is not simply a set of performance statistics, but a clear and urgent call to action. The immense potential of EBC as a true liquid biopsy for the lung will remain unrealized until the scientific community commits to a monumental, collaborative effort to standardize every step of the process. Its successful integration into clinical practice could dramatically improve the efficiency of lung cancer screening and ultimately, improve survival outcomes for patients with the world's most lethal malignancy.

6. References

1. Pérez-Sánchez C, Barbarroja N, Pantaleão LC, López-Sánchez LM, Ozanne SE, Jurado-Gámez B, et al. Clinical utility of microRNAs in exhaled breath condensate as biomarkers for lung cancer. J Pers Med. 2021; 11(2): 111.
2. Kazeminasab S, Ghanbari R, Emamalizadeh B, Jouyban-Gharamaleki V, Taghizadieh A, Jouyban A, et al. Exhaled breath condensate efficacy to identify mutations in patients with lung cancer: a pilot study. Nucleosides Nucleotides Nucleic Acids. 2022; 41(4): 370–83.
3. Tetik Vardarlı A, Ozgur S, Goksel T, Korba K, Karakus HS, Asık A, et al. Conversion of specific lncRNAs to biomarkers in exhaled breath condensate samples of patients with advanced stage non-small-cell lung cancer. Front Genet. 2023; 14: 1200262.
4. Santone D, Davis MD, Al Nasrallah N. Evaluating hydrogen peroxide in exhaled breath condensates for lung cancer detection. Am J Respir Crit Care Med. 2025; 211: A4816.
5. Tetik Vardarli A, Pelit L, Aldag C, Korba K, Celebi C, Dizdas TN, et al. Concordance in molecular genetic analysis of tumour tissue, plasma, and exhaled breath condensate

samples from lung cancer patients. J Breath Res. 2020; 14(3): 036001.
6. Rai D, Pattnaik B, Bangaru S, Bhatraju NK, Tak J, Kashyap S, et al. MicroRNAs in exhaled breath condensate: a pilot study of biomarker detection for lung cancer. Cancer Treat Res Commun. 2023; 35(100689): 100689.
7. Bano A, Yadav P, Sharma M, Verma D, Vats R, Chaudhry D, et al. Extraction and characterization of exosomes from the exhaled breath condensate and sputum of lung cancer patients and vulnerable tobacco consumers-potential noninvasive diagnostic biomarker source. J Breath Res. 2024; 18(4): 046003.
8. Carpagnano GE, Popov TA, Scioscia G, Ardò NP, Lacedonia D, Malerba M, et al. Exhaled breath temperature home monitoring to detect NSCLC relapse: Results from a pilot study. Biomed Res Int. 2022; 2022(1): 1515274.
9. Tetik Vardarlı A, Karakus HS, Korba K, Boluk B, Ozgur S, Gunduz C, et al. Assessment of exhaled breath condensate for ALK, RET, ROS1, and NTRK1 fusion transcript detection in NSCLC: Comparison with tissue and liquid biopsy samples. Thorac Cancer. 2025; 16(1): e15513.
10. Ahmed N, Kidane B, Wang L, Nugent Z, Moldovan N, McElrea A, et al. Metabolic alterations in sputum and exhaled breath condensate of early stage non-small cell lung cancer patients after surgical resection: a pilot study. Front Oncol. 2022; 12: 874964.
11. Kort S, Brusse-Keizer M, Schouwink H, Citgez E, de Jongh FH, van Putten JWG, et al. Diagnosing non-small cell lung cancer by exhaled breath profiling using an electronic nose: a multicenter validation study. Chest. 2023; 163(3): 697–706.
12. Hughes RT, Reno KE, Levine B, Bishop AC, Willey J, Farris M. A pilot study to detect inflammatory biomarkers of pneumonitis in exhaled breath during chemoradiotherapy for stage III non-small cell lung cancer. Int J Radiat Oncol Biol Phys. 2024; 120(2): e370.
13. Larracy R, Phinyomark A, Scheme E. Infrared cavity ring-down spectroscopy for detecting non-small cell lung cancer in exhaled breath. J Breath Res. 2022; 16(2): 026008.
14. Chen J-L, Han H-N, Lv X-D, Ma H, Wu J-N, Chen J-R. Clinical value of exhaled breath condensate let-7 in non-small cell lung cancer. Int J Clin Exp Pathol. 2020; 13(2): 163–71.
15. Wu J, Chen J, Lv X, Yang Q, Yao S, Zhang D, et al. Clinical value of serum and exhaled breath condensate inflammatory factor IL-11 levels in non-small cell lung cancer: Clinical value of IL-11 in non-small cell lung cancer. Int J Biol Markers. 2021; 36(2): 64–76.
16. Kordiak J, Szemraj J, Grabska-Kobylecka I, Bialasiewicz P, Braun M, Kordek R, et al. Intratumor heterogeneity and tissue distribution of KRAS mutation in non-small cell lung cancer: implications for detection of mutated KRAS oncogene in exhaled breath condensate. J Cancer Res Clin Oncol. 2019; 145(1): 241–51.
17. Chen J, Chen J, Lv X, Yang Q, Yao S. Epidermal growth factor in exhaled breath condensate as diagnostic method for non-small cell lung cancer. Technol Cancer Res Treat. 2019; 18: 1533033819872271.
18. Xie H, Chen J, Lv X, Zhang L, Wu J, Ge X, et al. Clinical value of serum and exhaled breath condensate miR-186 and IL-1β levels in non-small cell lung cancer. Technol Cancer Res Treat. 2020; 19: 1533033820947490.
19. Goggin C, Nay MO, Ryan D, Toomey S, Hennessy B, Calvert P. Exhaled breath condensate confirming T790M mutation in EGFR-mutated non-small cell lung cancer. Curr Probl Cancer Case Rep. 2021; 4(100114): 100114.

20. Wang S, Chu H, Wang G, Zhang Z, Yin S, Lu J, et al. Feasibility of detecting non-small cell lung cancer using exhaled breath condensate metabolomics. J Breath Res. 2025; 19(2).