Artificial Intelligence–Assisted Liquid Biopsy Analysis in Lung Cancer: Advances in Multi-Omic Biomarker Integration and Precision Oncology

Ishan Gowda[1]

Received March 2, 2026
Accepted July 5, 2026
Electronic access August 15, 2026

Cancer incidence continues to rise globally, driven by age-related accumulation of somatic mutations, with lung cancer remaining a leading cause of cancer-related mortality. Conventional tissue biopsies are invasive and limited in their ability to capture tumor heterogeneity and temporal evolution. Liquid biopsies, which analyze circulating tumor-derived biomarkers such as circulating tumor DNA (ctDNA), cell-free DNA (cfDNA), methylation patterns, and fragmentomic features, offer a minimally invasive alternative for longitudinal monitoring. When integrated with next-generation sequencing (NGS), these approaches generate large, complex datasets that require advanced novel methods. This review synthesizes recent advances (2020–2025) in artificial intelligence (AI)–assisted analysis of liquid biopsy data in lung cancer, with a focus on non-small cell lung cancer (NSCLC). This review systematically evaluates studies employing machine learning (ML) and deep learning (DL) techniques, including random forests, support vector machines, convolutional neural networks, and autoencoders. These models integrate multi-omic biomarkers such as ctDNA mutations, cfDNA methylation signatures, fragmentomics, and copy number alterations. Across studies, AI-enabled liquid biopsy analyses demonstrate improved sensitivity, specificity, and prognostic stratification compared to conventional analytical approaches. Deep learning methods show particular strength in high-resolution methylation and fragmentomic profiling, while classical ML models remain effective for survival prediction and risk stratification. Despite challenges related to cohort size, data bias, and ethical considerations, integrating AI with liquid biopsy technologies represents a transformative step toward precision oncology in lung cancer, enabling non-invasive early detection, personalized treatment, and minimal residual disease (MRD) monitoring.

Keywords: Tissue Biopsy, Liquid Biopsy, Next-Generation Sequencing, Multi-Omics, Lung Cancer, Artificial Intelligence.

Introduction

Cancer persists as the most prevalent and obscure disease in current medicine[1]. It is estimated that in the United States, 618,120 people will die from cancer in 2025, corresponding to approximately 1700 deaths every day[2]. Although medicine in the field of oncology has become more efficient and accurate in diagnosis, the rate of incidence is growing due to somatic mutations developing over a person's lifetime relative to age. Germline mutations occur within the deoxyribonucleic acid (DNA) in reproductive cells and are inheritable, passing mutations down to offspring. In contrast, somatic mutations occur in non-reproductive cells and are restricted to definite tissues. Approximately 90% of cancer patients develop cancer through somatic mutations[3]. Cancer is linked to both somatic and germline mutations; however, the increase in cancer cases in the past 10 years is attributed to increased age and the accumulation of somatic mutations over time[1]. Data collected from the World Health Organization [Figure 1] demonstrates the cumulative risk by age (in %) in the top 7 most prevalent cancers for males and females: pancreatic cancer, lung cancer, colorectal cancer, melanoma of skin, prostate cancer, kidney cancer, and breast cancer. Data were collected from the United States of America (US) and from people aged between 30-84 years. Data were gathered from the past 39 years and specifically from the years 1978-2017.

The graph represents a strong positive correlation between the risk of developing cancer and age in both males and females. Regarding males [Figure 1A], there is a linear growth in the risk of developing cancer till the age of 50 years. However, from ages 51–84, there is an exponential growth in the risk of cancer. Furthermore, the increased risk of developing pancreatic cancer, kidney cancer, and melanoma cancer from the ages of 51–84 is relatively low. There is approximately an increase of $\sim$1.8%-2.3% risk of developing pancreatic cancer, kidney cancer, and melanoma cancer with age. However, in males, the rate of developing colorectal cancer and lung

[1] *Cambridge Centre for International Research*

© The National High School Journal of Science 2026

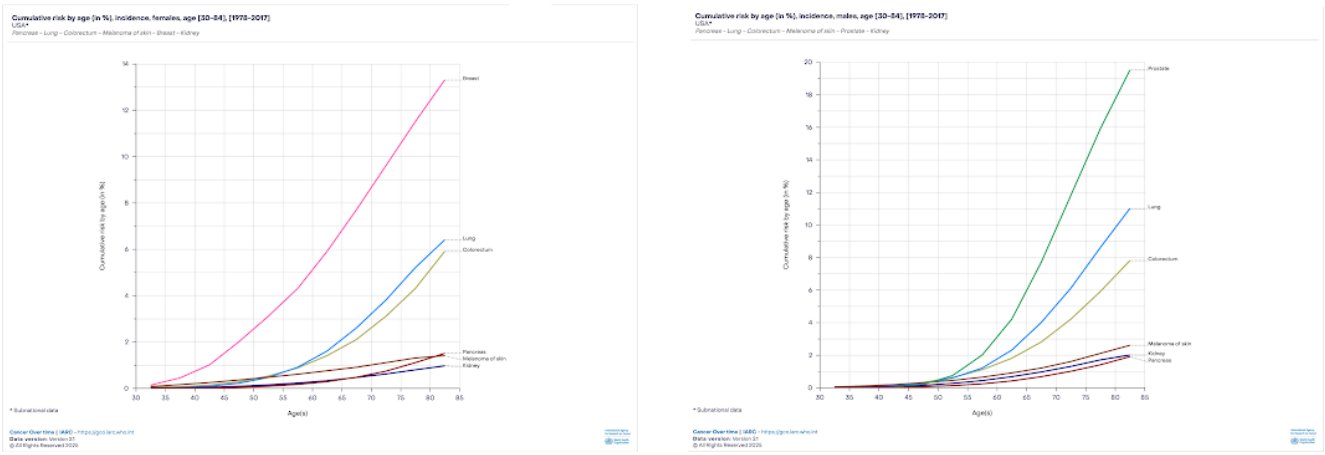

Figure 1. Correlation of age and risk for cancer development in males (A) and females (B) using the 6 highest risk cancer types for each sex in the USA.

cancer is relatively high between the ages of 55–84. There is approximately an increase of ∼7.5–10.5% risk of developing colorectal cancer and lung cancer between the ages of 55-84. Similar data were collected from the World Health Organization (GLOBALCAN), but for females [Figure 1B]. The graph represents a similar increase in risk between the ages of 51-84 in kidney cancer, melanoma skin, and pancreatic cancer. Additionally, similar to [Figure 1A], colorectal cancer and lung cancer risk have a positive correlation between the ages of 51-84, increasing in risk of developing cancer approximately ∼5.5%–6% with age.

In summary, [Figure 1A] and [Figure 1B] both show the highest increase in the risk of developing cancers with age, separated by gender. In [Figure 1A], males have the highest rate of developing prostate cancer at the age of 84, which is approximately 17%, and the rate of developing prostate cancer at the age of 50 is approximately 0.8%. This shows an approximately ∼16.2% increase from the age of 50. Similarly, females have the highest rate of developing breast cancer at the age of 84, which is approximately 13.4%, and the rate of developing breast cancer at the age of 50 is approximately 2.5%. This shows an approximately 10.9% increase from the age of 50. In conclusion, the data strongly support the increase in cumulative risk of developing cancer as age increases in both males and females. Although mutations are multifactoral due to risk contributors such as lifetime exposures, smoking, obesity, and immune changes, generally males and females older than 50 years are more prone to developing cancer due to somatic mutations developing in the body.

Method

For this literature review, PubMed and Google Scholar were used as the primary research engines. The literature search was conducted between October 2025 and March 2026 to identify relevant studies within the defined publication window. The keywords used included: "(Cancer) AND (Tumor naive assays), (Cancer) AND (Tumor informed assays), (Cancer) AND (Tissue biopsy), (Cancer) AND (Liquid biopsy), (Cancer) AND (ctDNA), (Cancer) AND (Next-generation sequencing), (Cancer) AND (Multi-cancer early detection), (Cancer) AND (Multi-omic Data), (Cancer) AND (Artificial intelligence), (Lung Cancer) AND (Artificial intelligence), (Lung Cancer) AND (Machine Learning), (Cancer) AND (Treatment methods), (Cancer) AND (AI ethics)". The exclusion criteria are as follows: books, letters, pre-prints, non-human studies, non-scholarly sources, and articles in languages other than English. The inclusion criteria are as follows: original research papers, literature review articles, clinical trials, clinical reports, and data from the years 2020–2025. A total of 295 articles were screened, of which 60 met the inclusion criteria. Studies were first screened based on title and abstract relevance, followed by full-text assessment for eligibility according to the inclusion and exclusion criteria. The methodological quality of included studies was evaluated based on the study design, relevance to the review topic, and clarity of reported methods. Studies were considered to meet acceptable quality standards if they clearly described their methodology, used validated analytical approaches, and reported clinically relevant outcomes in lung cancer or NSCLC contexts. Studies meeting the acceptable quality standards were included in the literature review.

Understanding lung cancer development

The cellular and molecular attributes of cancer, often described as the "Hallmarks of cancer", support the classifi-

© The National High School Journal of Science 2026

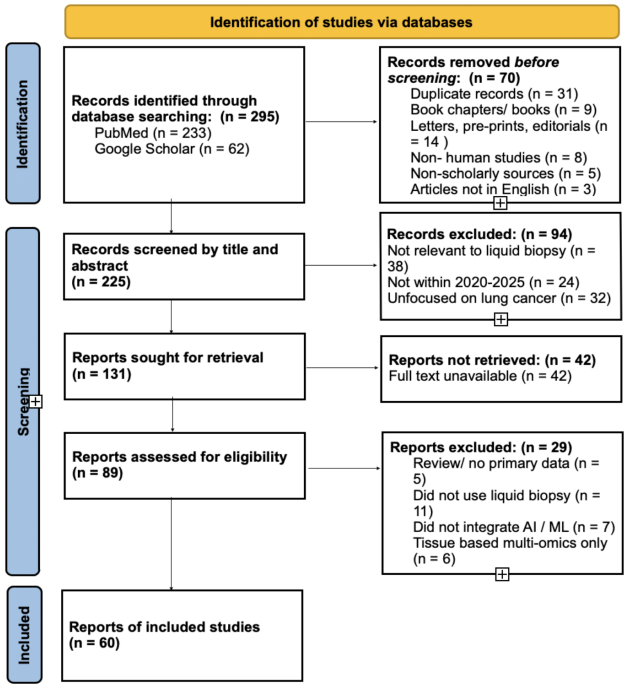

Figure 2. PRISMA flow diagram summarizing the study selection process for this review.

cation of therapeutic and tumor behavior strategies[4]. These hallmarks of cancer encompass the important biological characteristics tumor cells contain during cancer development, including sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, and activating invasion and metastasis[5]. Four newer hallmarks include unlocking phenotypic plasticity, nonmutational epigenetic reprogramming, polymorphic microbiomes, and senescent cells[5]. Integrating multi-omics approaches, which include genomics, transcriptomics, proteomics, metabolomics, and epigenomics, has become necessary when deciphering hallmarks, offering valuable insights into tumor heterogeneity[5,6].

Cancer is characterized as very heterogeneous. Cancer is composed not only of one type of cell; rather, it is composed of numerous types of cells throughout the body. Ultimately, this characteristic of cancer makes it difficult to diagnose, as it is highly variable in its cellular composition and molecular features, showing its importance in deciding treatment and diagnosis. Cancer exhibits multiple forms of mutations, including insertions, deletions, duplications, inversions, copy number variations (CNVs), and point mutations. Cancer is caused by specific mutations that are abundant in two gene categories, tumor suppressor genes and oncogenes. Somatic point mutations are abundant in most cancer types, such as the KRAS gene in colorectal cancer and the TP53 mutations in ovarian cancer[5]. The proto-oncogene KRAS is responsible for cell growth and division[5]. The oncogene KRAS can cause uncontrolled growth and tumor cell proliferation. The TP53 gene is responsible for initiating apoptosis, and a mutation causes an increased survival rate of mutated cells, causing neoplasms in the body[5]. In addition, germline mutations can also lead to cancer development. One example is the Hereditary Breast and Ovarian Cancer Syndrome (BRCA1/BRCA2), which is characterized by mutations in BRCA1 and BRCA2 genes, responsible for repairing double-stranded breaks in DNA[7].

Novel sequencing technologies in lung cancer diagnosis

Advances in technology have led to the transition from First Generation Sequencing (FGS), Sanger sequencing, and Maxam-Gilbert sequencing to Next Generation Sequencing (NGS). The Sanger and Maxam-Gilbert method of sequencing relies on the amplification of specific regions in the DNA and are followed by gel electrophoresis, which has a disadvantage in both cost and time compared to NGS[8]. Firstly, the Maxam-Gilbert method of sequencing has not been used recently due to its major disadvantages in the laboratory[8]. The high toxicity sourced from the phosphate isotope, difficulties analysing sequences longer than 500 base pairs (bp), and vast types of errors during cleavage all cause inefficiency during sequencing[8,9]. On the other hand, the Sanger method works by the use of dideoxynucleotides (ddNTPs)[10]. The ddNTPs are fluorescently labeled and can be read by a DNA sequencer machine that uses electrophoresis[8,10,11]. The ddNTPs are responsible for terminating strand elongation in the DNA so the machine can read the base pairs effectively[8,11]. Although the Sanger sequencing method has been proven to be a less toxic method than the Maxam-Gilbert method, it is slow and does not allow for numerous base pairs to be read in a short amount of time[8]. Using these methods to analyze a wide variety of regions will require excessive money and time, which is proven to be an impractical method[8,10]. Both methods are limited to a specific genetic locus, which is not beneficial when diagnosing cancer at an early stage, causing a need for a novel method[8,10].

To overcome the disadvantages of FGS, NGS is an improved method of sequencing that is more efficient and high-throughput in diagnosing diseases. NGS has allowed for the sequencing of the entire genome in a high-throughput manner, whole-genome Sequencing (WGS), and the entire exome by whole-exome sequencing (WES). NGS is comprised of many types of methods that are new and updated, which include Applied Biosystems (ABI) sequencing, Illumina sequencing, and Roche sequencing. Although each type of NGS method is

© The National High School Journal of Science 2026

used to sequence DNA, each method uses different methods to do so. NGS leads precision medicine to take a leap forward in diagnosing cancer at an early stage. Next-generation sequencing (NGS) has allowed for the specific profiling of epigenetic and genetic changes in the body, guiding targeted therapies[4]. Integrating genomic and epigenomic information by using NGS will allow for more precise prognoses and personalized treatment for each individual, understanding the genomic landscape and improving cancer diagnosis outcomes in many cancer types, known as precision medicine[4].

Lung cancer diagnosis using tissue biopsy

Tissue biopsy provides a conventional method used for cancer diagnosis. Tissue biopsies are a fundamental and established method for cancer diagnosis. The method enables molecular characterization of tumor tissue through sequencing and histopathological assessment. Tissue biopsies remain essential for definitive diagnosis and treatment planning in oncology. Tissue biopsies are composed of many different techniques specific to each location of cancer. The 5 most common tissue biopsy techniques used are endoscopic biopsies, incisional biopsies (IB), excisional biopsies, core needle biopsies (CNB), and fine needle aspiration biopsies (FNA). Endoscopic biopsies use jaw biopsy forceps inserted through the organ or a cavity to obtain a sample of tissue used for diagnosing and sequencing[12]. Endoscopic biopsies are commonly used for gastrointestinal cancers, lung cancers, bladder cancer, and cervical cancers[12]. Excisional biopsies are a process where a lesion is removed in its entirety, suspected to be cancerous[13]. Excisional biopsies are commonly used for oral cancers, melanomas, and deep suspicious processes[13]. IBs offer a large volume of tissue to be lesioned, while still leaving part of the tissue in the body[14]. IBs are used for sampling and allow for precise control over the incisional tract: near vessels and nerves[14]. However, the use of IB causes higher-risk complications that include hematoma[14]. IBs are most commonly used in almost all types of cancers[14]. FNA uses ultrasound to evaluate nodules suspected of malignancy[15]. FNA uses a thin needle to collect cells and/or fluids from a suspicious site[15]. Although FNA is a relatively non-invasive method for a tissue biopsy, it still has limitations. FNA is not reliable for several diagnostic categories, which include Bethesda I, Bethesda III, and Bethesda IV[16]. FNA is commonly used in thyroid cancers, breast cancers, lymph nodes, and lung cancer[15]. Finally, CNB uses a larger needle than FNA, the removal of a large, cylindrical structure of tissue that offers a better understanding of the heterogeneity of the tissue, which reduces false negatives[16]. However, CNB commonly creates scars at the site of extraction due to the needle being 6–10 cm in length, usually[16]. Typically, to obtain a sample using CNB, local anesthesia and a surgical technique are needed, which takes time and money[16]. CNB is commonly used in liver cancer, breast cancer, prostate cancer, kidney cancer, and sarcomas[16].

Tissue biopsies are an invasive method and have challenges that limit their success: time and spatial sampling. Tissue processing and sample collection take time to obtain, which is crucial when diagnosing baseline disease, MRD, and monitoring prognosis during treatment[17]. Moreover, tissue biopsies lack a spatial sample that has an accurate representation of the tumor in the body[18]. Due to spatial tumor heterogeneity and sampling limitations, tissue biopsies may lead to false-negative results or sampling bias[18]. This is because a single biopsy site may not fully capture the diversity of tumor subclones[18]. Additionally, tissue biopsies only sample the tumor from the time of the lesion; this does not allow for longitudinal monitoring[18]. Rather than replacing the method of tissue biopsy, a liquid biopsy is recognized as a complementary approach, which provides longitudinal and minimally invasive sampling[18].

Liquid biopsies as a noninvasive approach in lung cancer

Conversely, liquid Biopsy is a noninvasive method of collecting genetic material through body fluids, rather than collecting a sample from the body directly. Liquid biopsies are used for the detection of cancer-associated biomarkers across multiple cancer types, including lung cancer. The method shows a promising potential for early detection; however, performance in early-stage disease remains uncertain and is still under active research. Tumors have a characteristic of secretion, which is when genetic material is released into parts of the body when a cell enters apoptosis. The genetic material can be found in various forms, including circulating tumor cells (CTCs), cell-free DNA (cfDNA), and circulating tumor DNA (ctDNA). Specifically, this genetic material can be found in various locations, such as in plasma, lymphatic fluid, urine, and saliva. Once a sample is collected in a blood collection tube (BCT), it will be centrifuged to separate the blood into plasma. This sample is then used in an NGS machine to locate biomarkers exclusive to the cancer. However, liquid biopsies are not an alternative to tissue biopsies; rather, they serve as a complementary tool to increase the success of diagnosis[19]. Tissue biopsy remains the diagnostic standard. However, its invasive nature and spatial and temporal limitations make it a less suitable option for diagnosis, repeated sampling, and minimal residual disease (MRD) monitoring. More favorably, liquid biopsies are used to detect a wider range of cancer types and successfully diagnose at an earlier stage[19].

Specifically, ctDNA provides a noninvasive method for monitoring tumor burden by assessing baseline disease status and MRD, thereby supporting treatment decision-making[20].

© The National High School Journal of Science 2026

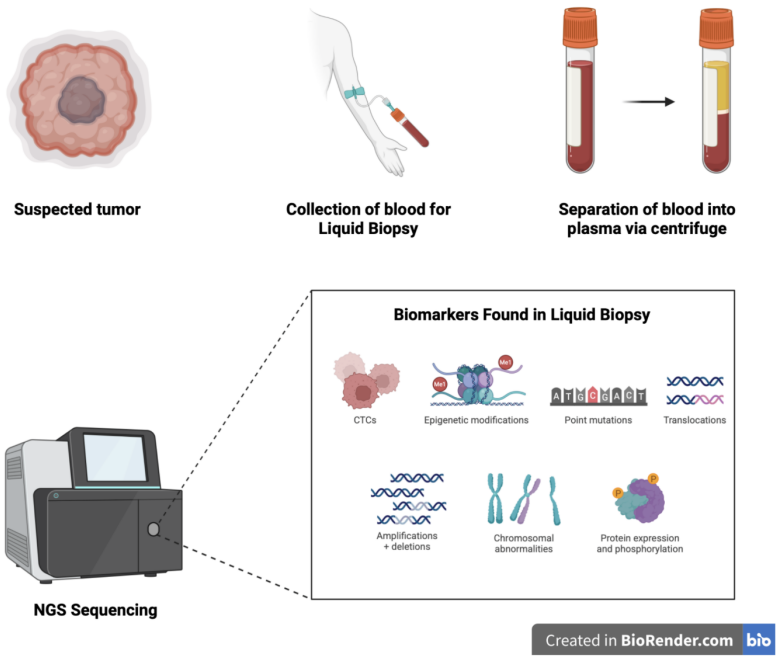

Figure 3. The process of liquid biopsy, showing how plasma is collected and isolated, and how it is analysed through sequencing for cancer diagnosis using various cancer biomarkers. This figure was created using BioRender.

The increased detection of ctDNA through liquid biopsy has allowed the ability to facilitate precision medicine by profiling tumors[21]. The cfDNA is isolated from plasma, following library preparation and sequencing. Sequencing technology is a process by which polymerase adds nucleotides to create a copied strand of DNA. Most notably, Illumina sequencing is a platform in which polymerase adds fluorescently labeled nucleotides, giving a unique color corresponding with each nucleotide type. This allows data to be collected, which can be analyzed for mutations. There are currently two methods of assays used to detect mutations in the ctDNA biomarker: tumor-naïve assays and tumor-informed assays[20].

Clinical applications of liquid biopsy in lung cancer

The use of a liquid biopsy allows for multiple biomarkers to be analyzed from a single sample. The suspicion of a tumor can derive from multiple variables, which include physical symptoms of pain, weakness, discomfort, and others. However, a suspicion of tumor growth does not always have to be present to collect a sample for a liquid biopsy. Cancer has an increased risk in older males and females, so tests and screening are being used for the older population without the suspicion of a tumor, aiding in early diagnosis. With the use of multiple biomarkers from a liquid biopsy, the specificity increases for precision medicine. Tumor-informed assays, where tumor tissue is sequenced from a biopsy sample, allow for personalized detection of malignancies. Tumor-informed assays use prior knowledge of a mutation in the body. This prior knowledge is gained through an invasive biopsy of a tumor growth. The sample is then sequenced to identify somatic mutations unique to the specific tumor. Through a liquid biopsy, the collection of ctDNA will go through additional sequencing, resembling the original sample of the tumorous growth. One method, named Integration of Variant Reads (INVAR), analyzes hundreds of thousands of known genomic variants and compares them to the mutated ctDNA. Tumor-informed assays provide a highly sensitive approach, which is beneficial when diagnosing MRD. Both tumor-informed and tumor-naïve assays have distinct clinical advantages and limitations, and their use depends on factors such as tissue availability, turnaround time, intended clinical application, and required sensitivity. However, tumor-informed assays require previous sequencing of the tissue, using a tissue biopsy, which uses time and money for the diagnosis process. Additionally, tumors are highly heterogeneous, consisting of a diverse population of phenotypic characteristics, which makes it difficult to diagnose cancer during baseline treatment and can lead to inaccurate results[22].

Alternatively, tumor-naïve assays are an alternative test that omits the sequencing of the tumor tissue. Tumor-naïve assays are a promising test as they target alterations of all types of

Table 1 Comparison of Tumor-Informed and Tumor-Naïve Liquid Biopsy Assays in Lung Cancer Applications.

Feature	Tumor-Informed Assays	Tumor-Naïve Assays
Prior tissue requirement	Requires prior tissue biopsy and sequencing	Does not require prior tissue sequencing
Primary approach	Tracks patient-specific known mutations	Detects broad genomic or epigenetic alterations
Typical biomarkers	ctDNA mutations unique to the tumor	Somatic mutations, CNVs, methylation patterns
Turnaround time	Longer due to tissue sequencing workflow	Generally shorter
Sensitivity	High sensitivity for MRD detection	Variable sensitivity depending on assay design
Common applications	MRD monitoring, recurrence surveillance	Early detection, screening, broad molecular profiling
Strengths	Personalized detection and high specificity	Faster workflow and broader applicability
Limitations	Requires tissue availability and added cost	May have lower specificity for patient-specific variants
Best-fit clinical setting	Longitudinal monitoring after diagnosis	Situations where tissue is unavailable or rapid profiling is needed

mutations, including somatic mutations, CNVs, and methylation changes. Within tumor-naïve assays, there are two types of approaches. Firstly, Genomic variant-based detection (gMRD) detects mutations based on DNA, finding common alterations in 750 cancer genes[23]. Most commonly, gMRD tumor-naïve assays detect the TP53 and the KRAS mutation as they are in high abundance in all tumor types. On the other hand, Methylation-based detection (mMRD) detects DNA methylation patterns that contain characteristics of cancer. Methyl groups (CH3) are chemical tags that control the production of proteins, which is a process important in epigenetics and aids in a non-genetic approach to diagnosis[23]. In summary, tumor-naïve assays offer advantages in turnaround time and reduced dependence on prior tissue sequencing, while tumor-informed assays may provide greater sensitivity for MRD detection through patient-specific mutation profiling.

Multi-Cancer Early Detection (MCED) test for earlier lung cancer detection

MCED promises to detect more than one type of cancer using a single test at the early stages, when people are asymptomatic. Advances in multiomics have enabled the development of liquid biopsies using cfDNA[24]. Multi-cancer early detection (MCED) blood tests improve the efficiency of traditional screening because they analyze a larger number of gene sequences in a shorter amount of time[24]. Additionally, MCED provides screening without any risks associated with whole-body imaging[24]. MCED is a novel approach when dealing with early diagnosis and diagnosis of MRD. Using an MCED approach powered by liquid biopsies, the ability to diagnose at an early stage becomes dominant[23].

MCED tests refer to the biomarkers they use, which include ctDNA, to detect cancer. MCED tests also have the responsibility of diagnosing the type and subtype of cancer using biomarkers. MCED tests continue to grow and expand as technology advances in oncology to benefit early detection accuracy and efficiency[25]. Common MCED tests include Galleri (GRAIL), Cancer Guard (Exact Sciences), CancerSEEK, PanSeer, Digital evaluation of fragmentation for early interception (DELFI), and OneTest[25].

GRAIL is a company that developed the GALLERI test that analyzes cfDNA fragments in the bloodstream and looks for abnormal methylation patterns to diagnose cancer diseases[26]. Unlike GRAIL, Cancer Guard analyzes ctDNA and tumor-associated proteins rather than Methylation patterns, giving a different approach to cancer diagnosis[27]. CancerSEEK is a specific blood test that detects ctDNA of the ovary, liver, stomach, pancreas, esophagus, colorectum, and lung, providing a less holistic approach[28]. Similar to CancerSEEK, PanSeer is a specific blood test that detects ctDNA of the stomach, esophagus, colorectal, lung, and liver. Both OneTest and DELFI use AI technology to provide an advanced view of cancer risks[29,30]. OneTest uses protein-based biomarkers to test a person's risk of cancer in the next 12 months, and uses AI by comparing the sample data to a large database of other individuals to aid in diagnosing accuracy[29]. Similarly, DELFI uses AI to compare the patient's cfDNA sample with a large database of cfDNA profiles from other individuals, improving the accuracy of identifying the cancer's tissue of origin[30].

AI and analysis of multi-omics data

Multiomics provides an approach to discovery that uses multiple levels of biological data[31]. Multiomics combines data from genomic, transcriptomic, epigenetic, proteomic, exosome, and clinical data sources[31]. Genomics focuses on the

© The National High School Journal of Science 2026

Table 2 Comparison of multi-cancer early detection (MCED) tests, summarizing their biomarker approaches, validation studies, and key clinical performance outcomes.

MCED Test	Biomarkers	Validation	Key Result
Galleri (GRAIL)	cfDNA methylation	Multi-cohort studies (healthy + cancer)	High accuracy across many cancers; predicts tissue of origin
CancerSEEK	ctDNA mutations + proteins	Case-control + screening cohorts	Detects early, treatable cancers; helps locate origin
PanSeer	cfDNA methylation	Pre-diagnosis blood samples	Detects cancer years before clinical diagnosis
DELFI	cfDNA fragment patterns	Multi-cancer cohorts + ML analysis	Improves detection and tissue classification
Cancer Guard (Exact Sciences)	ctDNA + proteins	Clinical + observational studies	Supports multi-cancer detection and profiling
OneTest	Protein biomarkers + AI	Retrospective datasets	Estimates short-term cancer risk

structure, function, and evolution of coded information within the human body in relation to cancer development[31]. Transcriptomics studies the complete set of RNA transcripts that define the expression of genes[31]. Epigenomics examines changes in gene expression that are not attributable to alterations in DNA sequence, but instead result from epigenetic modifications such as histone modifications and methylation patterns[31]. Proteomics evaluates protein expression and interactions that influence cellular function[31]. Exosome analysis focuses on extracellular vesicles released by cells that carry molecular information relative to tumor activity, and this analysis is also linked to the environmental factors that contribute to cancer development[31]. Clinical data integrates patient-specific information such as imaging, past medical history, and treatment outcomes to support personalized treatment[31]. The importance of multi-omics allows researchers to have a more comprehensive outlook on molecular changes that contribute to normal development, cellular response, and cancer disease.

Artificial intelligence and machine learning models are used to combine multiomic datasets using several computational methods. One common approach is feature concatenation, which merges variables from different omic datasets into a single dataset for model training[31]. Another method is dimensionality reduction, including techniques such as principal component analysis (PCA) and autoencoders, which simplify large datasets while retaining important biological information[31]. Multimodal fusion integrates different data types, including genomic, methylation, proteomic, fragmentomic, and clinical data, to improve cancer classification and biomarker identification[31]. Several studies summarized in [Table 8] used these approaches to combine ctDNA mutation profiles, methylation patterns, and clinical variables for lung cancer classification, prognostic assessment, and minimal residual disease (MRD) detection.

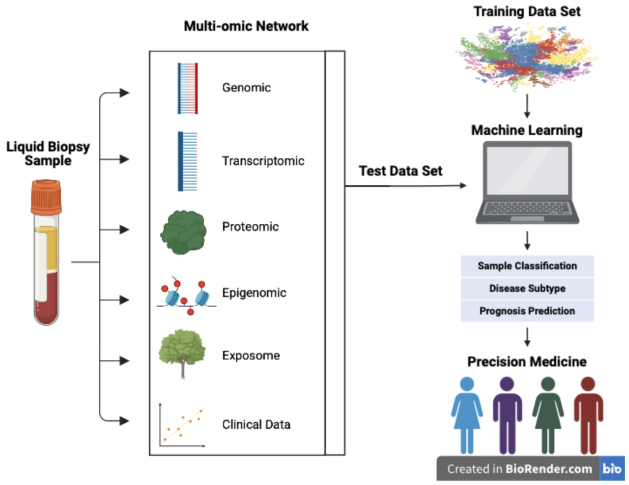

Figure 4. The process of multi-omic data from a liquid biopsy being used in machine learning for precision medicine. This figure was created using BioRender.

AI approach in cancer research: Lung cancer

This review summarizes the use of artificial intelligence (AI) in medicine, focusing on machine learning (ML) and deep learning (DL) approaches for cancer diagnosis. This review examines how AI analyzes liquid biopsy data to enhance detection and decision-making in lung cancer.

AI refers to a computer system specifically designed to perform tasks that require human intelligence. In medicine, AI supports decision-making and assists with the prediction of early diagnosis and prognosis monitoring[32]. AI is the broad category of human intelligence, which is divided into additional subsets that are more accurate at doing specific tasks. In the clinical setting, ML is "focused on algorithms that can

'learn' the patterns of training data and, subsequently, make accurate inferences about new data. This pattern recognition ability enables machine learning models to make decisions or predictions without explicit, hard-coded instructions"[33]. Similarly, DL is a "multilayered neural network whose design is inspired by the structure of the human brain"[33]. Unlike ML, DL uses a larger dataset to learn from, but in turn, produces more accurate decisions. In the clinical setting of diagnosis, ML and DL models are used in clinical studies. Within ML, Random Forest (RF) and Support Vector Machines (SVMs) are the most common in the clinical setting. Within DL, Convolutional Neural Networks (CNNs), Deep Neural Networks (DNNs), Recurrent Neural Networks (RNNs), and Autoencoders are among the most common.

These complex data sets include genomic sequences, electronic health records, and medical images[32]. In liquid biopsy studies, AI models do not directly identify the physical location of ctDNA within the body; rather, they analyze molecular and fragmentomic patterns to infer cancer classification, tissue-of-origin signatures, or disease progression. The integration of AI into oncology aims to enhance early diagnosis of cancer, improve accuracy, and optimize patient outcomes. Rather than replacing clinical providers, AI can complement their decisions and provide support to their choices. The role of artificial intelligence has greatly expanded in the field of oncology due to its ability to provide clinical decisions[32]. Although AI has a great influence, its coherence with medical providers remains unsettled. However, the analysis of recent studies can test the benefits of AI-incorporated liquid biopsy approaches in early diagnosis and MRD, aiming for increased accuracy[32].

Across recent studies between 2023-2025, ML and DL have become a core method in the analysis of liquid biopsy data for lung cancer studies. Incorporating ML has shown leverage on biomarker modalities and sequencing platforms to address diagnosis, prognosis, MRD, and treatment responses. The reviewed studies show supervised ML and DL approaches–including RF, CNN, and ensemble frameworks–in ctDNA mutations, cfDNA methylation patterns, fragmentomic features, and multimodal biomarker combinations. Cohorts in the studies ranged from small longitudinal monitoring studies to large, multi-validation datasets, with NGS techniques such as deep NGS, WGS, methylation-specific sequencing, and ultra-deep ctDNA profiling, supporting improved sensitivity and specificity in biomarker detection across multiple liquid biopsy modalities. Clinically, ML and

Table 3 The top 6 most utilized artificial intelligence (AI) models used in liquid biopsies that help in diagnosing lung cancer.

AI Technique	Typical Input Data	Analytical Task	Example Application in Liquid Biopsy Studies
Convolutional Neural Networks (CNNs)	cfDNA fragmentomics, methylation patterns, sequencing signal matrices	Pattern recognition and classification of high-dimensional molecular features	Used to classify lung cancer-associated fragmentomic and methylation signatures from cfDNA datasets[34].
Deep Neural Networks (DNNs)	Multiomic datasets including ctDNA mutations, methylation, proteomic, and clinical variables	Multimodal data integration and predictive modeling	Applied for cancer classification and prognostic prediction using integrated liquid biopsy datasets[35].
Random Forest (RF)	ctDNA mutation profiles, CNVs, methylation markers, clinical variables	Feature selection, biomarker ranking, and classification	Used to identify biomarker combinations associated with lung cancer diagnosis and prognosis[36].
Support Vector Machines (SVMs)	Multiomic biomarker panels and sequencing-derived features	Binary classification and risk stratification	Applied for distinguishing malignant versus nonmalignant molecular profiles in liquid biopsy samples[37].
Recurrent Neural Networks (RNNs)	Longitudinal ctDNA measurements and sequential clinical data	Temporal pattern analysis and disease monitoring	Used for tracking dynamic ctDNA changes during treatment response and MRD monitoring[38].
Autoencoders	High-dimensional sequencing and methylation datasets	Dimensionality reduction and latent feature extraction	Applied for identifying subtle molecular patterns and reducing noise in liquid biopsy datasets[39].

© The National High School Journal of Science 2026

Table 4 Overview of 9 recent studies (2023–2025) applying machine learning and deep learning methods to liquid biopsy data, including ctDNA, cfDNA, methylation patterns, fragmentomic features, and multimodal plasma biomarkers. Studies were selected based on predefined inclusion criteria requiring the use of liquid biopsy sampling and AI/ML modeling of blood-derived features. These studies represent distinct clinical application domains rather than a single unified predictive task. Studies included early detection, pulmonary nodule risk stratification, MRD monitoring, progression and survival prediction, and molecular subtype classification. Tumor tissue-based multi-omics studies were excluded to maintain consistency with liquid biopsy-derived data sources. The figure summarizes lung cancer type, AI/ML model, biomarker modality, cohort and validation, sequencing/data modality, modeling approach, and clinical utility (diagnosis, prognosis, or disease monitoring).

Author & Year	Lung Cancer Type	AI / ML Model	Biomarker Modality	Cohort & Validation	Sequencing / Data Modality	Modeling Approach	Clinical Utility	Performance
Shin et al., 2024 [40]	Lung cancer in COPD patients	Random Forest (RF) with cross-validation	ctDNA somatic mutations (VAF-based features)	n=177 COPD patients with newly diagnosed lung cancer; 30.5% with $\geq$1 detectable ctDNA mutation	Targeted deep NGS of predefined cancer genes	Supervised ML to associate clinical/imaging COPD features with ctDNA detectability	Identify factors influencing ctDNA shedding at diagnosis; feasibility of liquid biopsy in COPD	Not Available
Ding et al., 2024 [41]	Metastatic non-squamous NSCLC	Logistic regression; Cox proportional hazards; composite AUC ranking (4-fold CV)	ctDNA mutation profiles (e.g., TP53, KRAS, STK11, KEAP1)	Advanced NSCLC patients; median age 63; PD-L1 high in 22%; ABCP vs ACP treatment arms	Targeted plasma ctDNA sequencing	Longitudinal survival modeling (OS, PFS) using dynamic ctDNA features	Treatment response monitoring, progression risk stratification, and MRD assessment	AUC: 0.71–0.76 depending on ctDNA model; Sensitivity: 0.84 at 20% reduction cutoff
Kim et al., 2024 [42]	Mixed lung cancer types	CNN on 2-D methylation feature maps	cfDNA methylation markers (366 loci) + fragment size features	142 lung cancer vs 56 healthy controls; dilution series for LOD evaluation	WGEM-seq; MeDIP-seq; EM-seq; targeted methylation panels	Deep learning–based classification of methylation signatures	Non-invasive early lung cancer detection	AUC: 0.87; Sensitivity: Tumor fraction detection at 1% with 98% specificity; 0.1% at 80% specificity
Liang et al., 2024 [43]	Lung cancer with indeterminate pulmonary nodules	Multidimensional ML classifier	cfDNA methylation markers + clinical covariates (age, smoking)	n=963 participants; two independent cohorts (620 discovery / 343 validation); nodules 5–30 mm	Genome-wide and targeted cfDNA methylation NGS	Integrated molecular-clinical risk modeling	Early diagnosis and malignancy risk prediction for pulmonary nodules	AUC: 0.81 for methylation model; 0.90 internal and 0.89 external (cfDNA methylation + CT/risk factor model); Sensitivity: 0.88 internal and 0.90 external (methylation model); 0.97 internal and 0.95 external (combined model)

Jin et al., 2024 [44]	Lung cancer (screen-detected and clinical)	CNN-assisted imaging + multimodal ML	cfDNA methylation, ctDNA mutations, serum protein biomarkers	Training/validation 2:1; independent validation cohort n=291 (140 cancer, 25 benign, 126 controls)	Genome-wide methylation sequencing; UMI-based ultra-deep ctDNA sequencing	Multimodal feature fusion for classification	Early detection and post-screening nodule management	AUC: 0.953–0.966 (validation to independent validation; multi-omics screening model); Sensitivity: 99.2% overall (up to 100% in stage I–III)
Heeke et al., 2024 [45]	Small-cell lung cancer (SCLC-A/N/P/I)	Ensemble ML (500-model consensus); gene-ratio classifier	cfDNA methylation signatures; RNA-seq-derived subtype ratios	n=179 SCLC patients across two cohorts (105 discovery / 74 validation)	RRBS genome-wide methylation; tumor RNA-seq	Molecular subtype classification using epigenetic features	Non-invasive molecular subtyping of SCLC	AUC: 0.988 (SCLC detection in plasma cfDNA assay); Sensitivity: 100% (low cut-off, EpiScore = 65) / 94.0% (high cut-off)
Wu et al., 2024 [46]	Inoperable localized NSCLC	Survival SVM with LOOCV	Ultra-short cfDNA fragments ("neomers") + fragmentomics	n=44 patients; 97 longitudinal plasma samples	Targeted NGS + shallow WGS of plasma cfDNA	Time-to-event modeling using fragmentomic features	Disease monitoring and progression forecasting	Sensitivity: 60% (Neomer + Mutation, best combined model) / 45.5% (Neomer only) / 45% (Mutation only); Specificity: 92.3% (Neomer + Mutation) / 92.9% (Neomer) / 100% (Mutation, subset)
Widman et al., 2024 [47]	Advanced NSCLC (multi-cancer cohort)	Deep learning with ctDNA-specific feature engineering	WGS-based ctDNA SNVs & CNVs	Multi-cancer training/validation cohorts; lung cancer subset n≈22	Whole-genome cfDNA sequencing with ML-based noise suppression	Signal enrichment and longitudinal modeling	MRD detection and tumor burden monitoring	Not Available
Xu et al., 2024 [48]	Advanced NSCLC	ML-derived blood-based genomic immune subtypes (bGIS)	ctDNA genomic and immune-related features	n=460 patients from CHOICE-01 phase III trial (discovery + validation)	Deep targeted ctDNA NGS; low-pass WGS; tissue WES for concordance	Integrated genomic-immune subtype discovery	Prognostic stratification and therapy monitoring	Hazard-ratio: 0.66 (ctDNA+ vs ctDNA-, OS, chemo + PD-1 vs chemo) / 1.04 (ctDNA-, OS, no benefit) / 0.54 (bGIS-2 vs control, OS, chemo + PD-1 vs chemo) / ∼0.43–0.75 (bGIS-2 PFS benefit range across analyses)

© The National High School Journal of Science 2026

DL models have shown high utility in early lung cancer detection, malignancy risk stratification of pulmonary nodules, subtyping, MRD, survival prediction, and treatment response monitoring across non-small-cell and small-cell lung cancer subtypes.

In summary, these studies highlight that the use of ML and DL-aided liquid biopsy approaches can enhance diagnostic specificity and prognostic accuracy beyond the conventional methods. Moreover, the increased accuracy in clinical studies allows for a growing use of AI models in oncology, despite ongoing challenges related to generalizability, cohort heterogeneity, and clinical implementation.

Strengths and Limitations of Current AI-Assisted Liquid Biopsy Studies

Recent AI-assisted liquid biopsy studies show promising applications in lung cancer diagnosis, prognosis, MRD detection, and treatment monitoring. However, several limitations remain. One major strength across the reviewed studies is the increasing use of multimodal biomarker integration and independent validation cohorts. For example, Liang et al. evaluated 963 participants across separate discovery and validation cohorts, while Jin et al. incorporated an independent validation cohort for multimodal classification[43,44]. Similarly, Xu et al. analyzed 460 patients from the CHOICE-01 phase III clinical trial, improving the clinical relevance of prognostic modeling[48].

Despite these strengths, many studies remain limited by relatively small cohort sizes, restricted diversity, and limited prospective validation. For example, Wu et al. analyzed only 44 patients with 97 longitudinal plasma samples, which may reduce model generalizability across broader NSCLC populations[46]. Additionally, Widman et al. included a relatively small lung cancer subset within a larger multi-cancer cohort, which limits lung cancer-specific conclusions[47]. Lastly, several studies relied on retrospective cohort designs and internal cross-validation approaches, which may increase the risk of overfitting and data leakage. Class imbalance also remains an important challenge in AI-based liquid biopsy studies. Some studies included substantially different numbers of cancer and control samples. For example, Kim et al. evaluated 142 lung cancer patients and 56 healthy controls[42]. Imbalanced datasets may influence model sensitivity and specificity. This is more particular in early-stage disease detection, where ctDNA abundance is low.

Overall, current ML and DL models demonstrate strong potential for improving liquid biopsy analysis in lung cancer. However, larger cohort sizes, unrestricted diversity, and nonlimited prospective validations are necessary before large clinical implementation can occur.

Using AI towards personalized treatment

The use of biomarkers allows oncologists to tailor treatments to be used specifically for each patient's molecular profile[49,50]. Personalized treatment methods are used to treat baseline disease and also MRD. Conventional methods of treatment use chemotherapy, radiotherapy, and surgery. These treatment methods aim to kill fast-dividing and growing cells. While these are options for treating cancer, it is an invasive methods as it harms healthy growing cells, leading to complications and pain[49]. These conventional treatment methods have additional challenges, which necessitate the need for more personalized treatment.

The first challenge, resistance to targeted therapies, remains a core issue in targeted treatment[49]. Cancer develops acquired and primary resistance to treatment methods. This is in direct cause from co-mutations, alternative signaling pathways, and mutations emerging during the treatment process[49]. The second challenge is associated with multigene sequencing using NGS[49]. Although NGS is essential for precision medicine, daily access and implementation remain limited in clinical practice[49]. This is due to unequal cost, infrastructure, and a valid framework to interpret these genomic result outputs[49]. The third challenge is associated with drug development. Current drugs still heavily rely on organ-based randomized trials, which slow down the process for treating cancer[49]. This is due to the scarcity of cases for specific and rare cancer types and subtypes. For these rare cases, real-world data from cases is limited, so precision cancer treatment is still unbalanced[49]. Additionally, this is a flaw with ML and DL incorporation into liquid biopsies and treatment analysis, due to the lack of data to make decisions. The final challenge is associated with tumor heterogeneity. Because tumors consist of many cells and profiles, it is difficult to treat the whole tumorous growth[6,49]. It is important to note that tumor heterogeneity does not only lie in genomics; however, it consists of multiomics, which shows why multiomic data is important for diagnosis and deciding treatment methods[5,6,49]. The use of spatial biological profiling and ex vivo models needs to be used to treat these cancers to gain a full genetic map of the tumor[5,6,49].

To overcome these challenges, advanced treatment options are typically used. These novel treatment options include hormone therapy, targeted therapy, immunotherapy, and stem-cell therapy. Hormone therapies block specific hormones that aid in cancer growth and proliferation, specifically in the cancer, allowing for fewer complications[51]. Similarly, targeted therapy uses drugs, most commonly monoclonal antibodies, that help block molecular properties that help in proliferation[52]. Immunotherapy uses a slower method for fighting cancer. Immunotherapy aids in boosting the immune system to fight the cancer itself, allowing for a noninvasive method[53,54]. Lastly, stem cell therapy uses lab-engineered cells that are built to

attack proliferating cancer cells in the body[55]. In conclusion, liquid biopsies are needed to identify the properties of the cancer, which guide advanced treatment options. With the emerging use of ML and DL incorporated into liquid biopsies, cancer's multiomic profile will be better understood, providing increased personalized medicine[49,50,56].

Ethical considerations using AI and liquid biopsy

When using AI, ethical considerations should be addressed. On average, people report that they are comfortable with the use of AI in unrelated healthcare tasks[57]. These unrelated healthcare tasks include scheduling appointments, follow-ups, and meetings, with 84% of respondents comfortable with this limited use of AI[57]. However, when respondents were asked about their comfort with AI use in healthcare, "loss of human touch" was a common theme among respondents[57]. Respondents reported that AI will lose human decision-making control, as they were less comfortable with AI decision-making programs[57,58]. Moreover, a small percentage of respondents say AI-incorporated decision-making is the right step in healthcare; however, oversight should be implemented by doctors, and over-reliance on AI by doctors should not be met[57,58].

Additionally, data collection for AI training has ethical considerations. Patient data, including genomic and clinical data, needs to remain confidential and be obtained with patient consent, which is difficult as the data set needed for AI models to output accurate decisions is relatively large[59]. Additionally, patients want transparency with what their patient data is being used for and how it is being implemented, which takes time and money[59].

Moreover, bias considerations are important when considering AI-incorporated genomics and pharmacogenomics. Although data collection is improving, current genomic databases train their AI models disproportionately, as they are derived from a majority European population[60]. This creates a diversity gap, which favors the treatment of European people rather than minorities[60]. This lack of heterogeneous data results in biased predictions and reduced clinical accuracy for the underrepresented racial groups[59]. Moreover, this biased data has a direct genetic influence on how patients metabolize medications in cancer treatment[60]. For example, clopidogrel (Plavix), a commonly prescribed antiplatelet drug, is metabolized by the CYP2C19 enzyme[60]. Variants in the CYP2C19 gene are very common across racial groups and can lead to increased or decreased drug activation of cancer medication.

Table 5 Strengths and Limitations of AI-Based Liquid Biopsy Studies in Lung Cancer.

Study	Validation Strategy	Major Strength	Key Limitation
Shin et al., 2024[40]	Internal validation with training/testing cohorts	Integration of ctDNA with clinical COPD features	Limited cohort size and single disease population
Ding et al., 2024[41]	4-fold cross-validation	Longitudinal ctDNA survival modeling	Limited external validation reported
Kim et al., 2024[42]	Internal validation and dilution testing	Deep learning methylation and fragmentomic profiling	Potential class imbalance between cancer and control groups
Liang et al., 2024[43]	Independent discovery and validation cohorts	Large multi-cohort validation design	Prospective clinical implementation has not yet been established
Jin et al., 2024[44]	Independent validation cohort	Multimodal biomarker integration	Complex multimodal workflow may limit scalability
Heeke et al., 2024[45]	Cross-cohort methylation validation	Epigenetic subtype classification	Limited information regarding external institutional validation
Wu et al., 2024[46]	Internal validation with longitudinal cohort analysis	Longitudinal fragmentomic monitoring	Small sample size may limit generalizability
Widman et al., 2024[47]	Multi-cohort deep learning validation	Whole-genome ctDNA signal enrichment	Small lung cancer subgroup size
Xu et al., 2024[48]	Discovery and validation cohorts from phase III trial	Clinical trial-associated genomic modeling	Primarily focused on advanced-stage disease

© The National High School Journal of Science 2026

This affects treatment efficacy and patient safety due to mistreatment[60]. The variant frequency differs substantially across ethnic populations, meaning that AI-driven clinical tools may fail to accurately predict drug responses in diverse populations, due to the homogeneous data given[59]. Including more heterogeneous data could offset the biased data and lead to improved reliability of AI-guided therapy and diagnosis[59,60].

Despite promising results, this review is limited by the heterogeneity of study cohorts and variations in AI/ML model implementation, which may introduce bias. Additionally, the studies primarily reflect populations from specific regions, including East Asia, the United States, and the United Kingdom, and genetic differences across these populations may limit the generalizability of the findings.

Conclusion

Overall, it is clear that there is a positive correlation between aging and somatic mutations, which is linked to higher incidence rates. To diagnose cancers, traditional tissue biopsies provide an invasive method of sampling. Alternatively, liquid biopsies are a complementary tool to tissue biopsies, which provide a non-invasive approach to diagnosis and offer real-time monitoring. The use of liquid biopsies and NGS allows for the targeting and detection of a wide range of biomarkers in cfDNA. With advanced technologies, the field is moving towards multiomics, including genomics, epigenetics, proteomics, metabolomics, and fragmentomics. Additionally, using ML approaches, liquid biopsy has shown positive results in lung cancer specificity in a large number of patients. Still, standardization of protocols and analytical tools among private companies and public hospitals needs to be addressed before it is implemented into routine clinical care. Advances in liquid biopsies allow for precise and targeted treatments that drive personalized medicine.

References

1 H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. Vol. 71, pg. 209-249, 2021, https://doi.org/10.3322/caac.21660.

2 R. L. Siegel, T. B. Kratzer, A. N. Giaquinto, H. Sung, A. Jemal. Cancer statistics, 2025. CA: A Cancer Journal for Clinicians. Vol. 75, pg. 10-45, 2025, https://doi.org/10.3322/caac.21871.

3 S. Darabi, J. P. Homer, C. B. Torres, C. E. Zuazo, V. Dalili-Shoaie, M. J. Demeure, D. R. Braxton. Somatic profiling prospectively guides germline genetic testing in patients with mutations in high-risk cancer predisposition genes. Oncology Issues. Vol. 39, pg. 67-75, 2024, https://doi.org/10.3928/25731777-20240806-08.

4 S. N. Gibbs, D. Peneva, G. C. Cuyun Carter, M. R. Palomares, S. Thakkar, D. W. Hall, H. Dalglish, C. Campos, I. Yermilov. Comprehensive review on the clinical impact of next-generation sequencing tests for the management of advanced cancer. JCO Precision Oncology. Vol. 7, pg. –, 2023, https://doi.org/10.1200/po.22.00715.

5 D. Hanahan, R. Weinberg. Hallmarks of cancer: The next generation. Cell. Vol. 144, pg. 646-674, 2011, https://doi.org/10.1016/j.cell.2011.02.013.

6 M. Dong, L. Wang, N. Hu, Y. Rao, Z. Wang, Y. Zhang. Integration of multi-omics approaches in exploring intra-tumoral heterogeneity. Cancer Cell International. Vol. 25, pg. –, 2025, https://doi.org/10.1186/s12935-025-03944-2.

7 R. Roy, J. Chun, S. N. Powell. BRCA1 and BRCA2: Different roles in a common pathway of genome protection. Nature Reviews Cancer. Vol. 12, pg. 68-78, 2011, https://doi.org/10.1038/nrc3181.

8 K. Eren, N. Taktakoglu, I. Pirim. DNA sequencing methods: From past to present. The Eurasian Journal of Medicine. Vol. 54, pg. S47-S56, 2023, https://doi.org/10.5152/eurasianjmed.2022.22280.

9 L. T. França, E. Carrilho, T. B. Kist. A review of DNA sequencing techniques. Quarterly Reviews of Biophysics. Vol. 35, pg. 169-200, 2002, https://doi.org/10.1017/s0033583502003797.

10 W. R. McCombie, J. D. McPherson, E. R. Mardis. Next-generation sequencing technologies. Cold Spring Harbor Perspectives in Medicine. Vol. 9, pg. a036798, 2018, https://doi.org/10.1101/cshperspect.a036798.

11 C. M. Rubin, C. W. Schmid. Pyrimidine-specific chemical reactions useful for DNA sequencing. Nucleic Acids Research. Vol. 8, pg. 4613-4620, 1980, https://doi.org/10.1093/nar/8.20.4613.

12 Vikram Bhatia, Phool Chand, Sashi Kumar. Tissue Acquisition and Handling in Gastrointestinal Endoscopy. Tropical Gastroenterology. Vol. 44, pg. 4–15, 2023, https://doi.org/10.7869/tg.731.

13 M. B. Sequeira Campos, C. J. Beard, S. Ponnarasu, G. J. Schmieder. Excisional biopsy. StatPearls Publishing. 2025.

14 A. Klein, T. Fell, C. Birkenmaier, J. Fromm, V. Jansson, T. Knösel, H. R. Dürr. Relative sensitivity of core-needle biopsy and incisional biopsy in the diagnosis of musculoskeletal sarcomas. Cancers. Vol. 13, pg. 1393, 2021, https://doi.org/10.3390/cancers13061393.

15 I. Levine, A. J. Trindade. Endoscopic ultrasound fine needle aspiration fine needle biopsy for pancreatic masses, subepithelial lesions, and lymph nodes. World Journal of Gastroenterology. Vol. 27, pg. 4194-4207, 2021, https://doi.org/10.3748/wjg.v27.i26.4194.

16 D. D. Dolidze, S. Covantsev, G. M. Chechenin, N. V. Pichugina, A. V. Bedina, A. Bumbu. Core needle biopsy for thyroid nodules assessment—a new horizon? World Journal of Clinical Oncology. Vol. 15, pg. 580-586, 2024, https://doi.org/10.5306/wjco.v15.i5.580.

17 A. S.-Y. Leong, D. Price. Incorporation of microwave tissue processing into a routine pathology laboratory: Impact on turnaround times and laboratory work patterns. Pathology. Vol. 36, pg. 321-324, 2004, https://doi.org/10.1080/00313020410001721564.

18 L. Ma, H. Guo, Y. Zhao, Z. Liu, C. Wang, J. Bu, T. Sun, J. Wei. Liquid biopsy in cancer: current status, challenges and future prospects. Signal Transduction and Targeted Therapy. Vol. 9, pg. 336, 2024, https://doi.org/10.1038/s41392-024-02021-w.

19 Z. Qureshi, F. Altaf, M. Khanzada, A. Safi, Z. Asghar, D. Warraich, S. Shah. Liquid biopsies for early detection and monitoring of cancer: Advances, challenges, and future directions. Annals of Medicine Surgery. Vol. 87, pg. 3244-3253, 2025, https://doi.org/10.1097/ms9.0000000000002776.

20 O. Abidoye, D. H. Ahn, M. J. Borad, C. Wu, T. Bekaii-Saab, S. Chakrabarti, M. B. Sonbol. Circulating tumor DNA testing for minimal residual disease and its application in colorectal cancer. Cells. Vol. 14, pg. 161, 2025, https://doi.org/10.3390/cells14030161.

21 M. Nikanjam, S. Kato, R. Kurzrock. Liquid biopsy: Current technology and clinical applications. Journal of Hematology Oncology. Vol. 15, pg. –, 2022, https://doi.org/10.1186/s13045-022-01351-y.

22 J. Zhao, J. Reuther, K. Scozzaro, M. Hawley, E. Metzger, M. Emery, I.

Chen, M. Barbosa, L. Johnson, A. O'Connor, M. Washburn, L. Hartje, E. Reckase, V. Johnson, Y. Zhang, E. Westheimer, W. O'Callaghan, N. Malani, A. Chesh, R. Daber. Personalized cancer monitoring assay for the detection of ctDNA in patients with solid tumors. Molecular Diagnosis Therapy. Vol. 27, pg. 753-768, 2023, https://doi.org/10.1007/s40291-023-00670-1.

23 G. E. Alexander, W. Lin, F. E. Ortega, M. Ramaiah, B. Jung, L. Ji, E. Revenkova, P. Shah, C. Croisetiere, J. R. Berman, L. Eubank, G. Naik, J. Brooks, A. Mich, S. Shojaee, N. Ronaghi, H. Chawla, X. Hou, Q. Liu, N. Hunkapiller. Analytical validation of a multi-cancer early detection test with cancer signal origin using a cell-free DNA–based targeted methylation assay. PLOS ONE. Vol. 18, pg. e0283001, 2023, https://doi.org/10.1371/journal.pone.0283001.

24 D. Schrag, T. M. Beer, C. H. McDonnell 3rd, L. Nadauld, C. A. Dilaveri, R. Reid, C. R. Marinac, K. C. Chung, M. Lopatin, E. T. Fung, E. A. Klein. Blood-based tests for multicancer early detection (PATHFINDER): a prospective cohort study. Lancet. Vol. 402, pg. 1251-1260, 2023, https://doi.org/10.1016/S0140-6736(23)01700-2.

25 G. Wei, H. Jiang. Multi-cancer early detection tests: Pioneering a revolution in cancer screening. Clinical Cancer Bulletin. Vol. 3, pg. –, 2024, https://link.springer.com/article/10.1007/s44272-024-00015-x.

26 Klein EA, Richards D, Cohn A, et al. Clinical validation of a targeted methylation-based multi-cancer early detection test using an independent validation set. Annals of Oncology. Vol. 32, no. 9, pg. 1167-1177, 2021, https://doi.org/10.1016/j.annonc.2021.05.806.

27 Cohen JD, Li L, Wang Y, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science. Vol. 369, no. 6499, 2020, https://doi.org/10.1126/science.abb9601.

28 Lennon AM, Buchanan AH, Kinde I, et al. Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. Science. Vol. 369, no. 6499, 2020, pg. eabb9601, https://doi.org/10.1126/science.abb9601.

29 Heitzer E, Haque IS, Roberts CES, Speicher MR. Current and future perspectives of liquid biopsies in genomics-driven oncology. Nature Reviews Genetics. Vol. 24, pg. 592-606, 2023, https://doi.org/10.1038/s41576-023-00576-2.

30 Mathios D, Johansen JS, Cristiano S, et al. Detection and characterization of cancers using cell-free DNA fragmentomes. Nature Communications. Vol. 12, pg. 5060, 2021, https://doi.org/10.1038/s41467-021-24994-0.

31 Y. J. Heo, C. Hwa, G.-H. Lee, J.-M. Park, J.-Y. An. Integrative multi-omics approaches in cancer research: From biological networks to clinical subtypes. Molecules and Cells. Vol. 44, pg. 433-443, 2021, https://doi.org/10.14348/molcells.2021.0042.

32 L. Wei, D. Niraula, E. D. H. Gates, J. Fu, Y. Luo, M. J. Nyflot, S. R. Bowen, I. M. El Naqa, S. Cui. Artificial intelligence (AI) and machine learning (ML) in precision oncology: A review on enhancing discoverability through multiomics integration. The British Journal of Radiology. Vol. 96, pg. –, 2023, https://doi.org/10.1259/bjr.20230211.

33 O. Colliot. A non-technical introduction to machine learning. Machine Learning for Brain Disorders, NCBI Bookshelf. 2023, https://www.ncbi.nlm.nih.gov/books/NBK597507/.

34 A. Derry, M. Krzywinski, N. Altman. Convolutional neural networks. Nature Methods. Vol. 20, pg. 1269-1270, 2023, https://doi.org/10.1038/s41592-023-01973-1.

35 Y. LeCun, Y. Bengio, G. Hinton. Deep learning. Nature. Vol. 521, pg. 436-444, 2015, https://doi.org/10.1038/nature14539.

36 L. Breiman. Random forests. Machine Learning. Vol. 45, pg. 5-32, 2001, https://doi.org/10.1023/a:1010933404324.

37 J. G. Greener, S. M. Kandathil, L. Moffat, D. T. Jones. A guide to machine learning for biologists. Nature Reviews Molecular Cell Biology. Vol. 23, pg. 40-55, 2021, https://doi.org/10.1038/s41580-021-00407-0.

38 S. Das, A. Tariq, T. Santos, S. S. Kantareddy, I. Banerjee. Recurrent neural networks (RNNs): Architectures, training tricks, and introduction to influential research. Neuromethods. pg. 117-138, 2023, https://doi.org/10.1007/978-1-0716-3195-9_4.

39 B.-W. Han, X. Yang, S.-F. Qu, Z.-W. Guo, L.-M. Huang, K. Li, G.-J. Ouyang, G.-X. Cai, W.-W. Xiao, R.-T. Weng, S. Xu, J. Huang, X.-X. Yang, Y.-S. Wu. A Deep-Learning Pipeline for TSS Coverage Imputation From Shallow Cell-Free DNA Sequencing. Frontiers in Medicine. Vol. 8, pg. 684238, 2021, https://doi.org/10.3389/fmed.2021.684238.

40 S. H. Shin, S. Cha, H. Y. Lee, S.-H. Shin, Y. J. Kim, D. Park, K. Y. Han, Y. J. Oh, W.-Y. Park, M.-J. Ahn, H. Kim, H. H. Won, H. Y. Park. Machine learning model for circulating tumor DNA detection in chronic obstructive pulmonary disease patients with lung cancer. Translational Lung Cancer Research. Vol. 13, pg. 112-125, 2024, https://doi.org/10.21037/tlcr-23-633.

41 H. Ding, M. Yuan, Y. Yang, X. S. Xu. Identifying key circulating tumor DNA parameters for predicting clinical outcomes in metastatic non-squamous non-small cell lung cancer after first-line chemoimmunotherapy. Nature Communications. Vol. 15, pg. –, 2024, https://doi.org/10.1038/s41467-024-51316-7.

42 M. Kim, J. Park, S. H. Jeong, B.-H. Byun, Y. Shin, Y. Im, J. Cho, E.-H. Cho. Deep learning model integrating cfDNA methylation and fragment size profiles for lung cancer diagnosis. Scientific Reports. Vol. 14, pg. –, 2024, https://doi.org/10.1038/s41598-024-63411-2.

43 W. Liang, J. Tao, C. Cheng, H. Sun, Z. Ye, S. Wu, Y. Guo, J. Zhang, Q. Chen, D. Liu, L. Liu, H. Tian, L. Teng, N. Zhong, J.-B. Fan, J. He. A clinically effective model based on cell-free DNA methylation and low-dose CT for risk stratification of pulmonary nodules. Cell Reports Medicine. Vol. 5, pg. 101750, 2024, https://doi.org/10.1016/j.xcrm.2024.101750.

44 Y. Jin, W. Mu, Y. Shi, Q. Qi, W. Wang, Y. He, X. Sun, B. Yang, P. Cui, C. Li, F. Liu, Y. Liu, G. Wang, J. Zhao, Y. Zhang, S. Zhang, C. Cao, C. Sun, N. Hong, K. Chen. Development and validation of an integrated system for lung cancer screening and post-screening pulmonary nodules management: A proof-of-concept study (ASCEND-LUNG). EClinicalMedicine. Vol. 75, pg. 102769, 2024, https://doi.org/10.1016/j.eclinm.2024.102769.

45 S. Heeke, C. M. Gay, M. R. Estecio, H. Tran, B. B. Morris, B. Zhang, X. Tang, M. G. Raso, P. Rocha, S. Lai, E. Arriola, P. Hofman, V. Hofman, P. Kopparapu, C. M. Lovly, K. Concannon, L. G. De Sousa, W. E. Lewis, K. Kondo, Y. Raviv. Tumor- and circulating-free DNA methylation identifies clinically relevant small cell lung cancer subtypes. Cancer Cell. Vol. 42, pg. 225-237.e5, 2024, https://doi.org/10.1016/j.ccell.2024.01.001.

46 Y. Wu, C. Li, Y. Yang, T. Zhang, J. Wang, W. Tang, N. Li, H. Bao, X. Wang, N. Bi. Predicting disease progression in inoperable localized patients using machine learning model. Cancer Medicine. Vol. 13, pg. –, 2024, https://doi.org/10.1002/cam4.70316.

47 A. J. Widman, M. Shah, A. Frydendahl, D. Halmos, C. C. Khamnei, N. Øgaard, S. Rajagopalan, A. Arora, A. Deshpande, W. F. Hooper, J. Quentin, J. Bass, M. Zhang, T. Langanay, L. Andersen, Z. Steinsnyder, W. Liao, M. H. Rasmussen, T. V. Henriksen, L. Spain. Ultrasensitive plasma-based monitoring of tumor burden using machine-learning-guided signal enrichment. Nature Medicine. Vol. 30, pg. 1655-1666, 2024,https://doi.org/10.1038/s41591-024-03040-4.

48 J. Xu, R. Wan, Y. Cai, S. Cai, L. Wu, B. Li, J. Duan, Y. Cheng, X. Li, X. Wang, L. Han, X. Wu, Y. Fan, Y. Yu, D. Lv, J. Shi, J. Huang, S. Zhou, B. Han, Y. Song. Circulating tumor DNA-based stratification strategy for chemotherapy plus PD-1 inhibitor in advanced non-small-cell lung cancer. Cancer Cell. Vol. 42, pg. 1598-1613.e4, 2024, https://doi.org/10.1016/j.ccell.2024.08.013.

49 A. Passaro, M. Al Bakir, E. G. Hamilton, M. Diehn, F. André, S. Roy-

© The National High School Journal of Science 2026

Chowdhuri, G. Mountzios, I. I. Wistuba, C. Swanton, S. Peters. Cancer biomarkers: Emerging trends and clinical implications for personalized treatment. Cell. Vol. 187, pg. 1617-1635, 2024, https://doi.org/10.1016/j.cell.2024.02.041.

50 L. Paleari. Personalized assessment for cancer prevention, detection, and treatment. International Journal of Molecular Sciences. Vol. 25, pg. 8140, 2024, https://doi.org/10.3390/ijms25158140.

51 T. Deli, M. Orosz, A. Jakab. Hormone replacement therapy in cancer survivors – review of the literature. Pathology Oncology Research. Vol. 26, pg. 63-78, 2019, https://doi.org/10.1007/s12253-018-00569-x.

52 Y. Liu, H. Wang. Biomarkers and targeted therapy for cancer stem cells. Trends in Pharmacological Sciences. Vol. 45, pg. 56-66, 2024, https://doi.org/10.1016/j.tips.2023.11.006.

53 R. Rui, L. Zhou, S. He. Cancer immunotherapies: Advances and bottlenecks. Frontiers in Immunology. Vol. 14, pg. –, 2023, https://doi.org/10.3389/fimmu.2023.1212476. G. L. Szeto, S. D. Finley. Integrative approaches to cancer immunotherapy. Trends in Cancer. Vol. 5, pg. 400-410, 2019, https://doi.org/10.1016/j.trecan.2019.05.010.

54 G. L. Szeto, S. D. Finley. Integrative approaches to cancer immunotherapy. Trends in Cancer. Vol. 5, pg. 400-410, 2019, https://doi.org/10.1016/j.trecan.2019.05.010.

55 L. Yang, P. Shi, G. Zhao, J. Xu, W. Peng, J. Zhang, G. Zhang, X. Wang, Z. Dong, F. Chen, H. Cui. Targeting cancer stem cell pathways for cancer therapy. Signal Transduction and Targeted Therapy. Vol. 5, pg. –, 2020, https://doi.org/10.1038/s41392-020-0110-5.

56 F. Mosele, E. Deluche, A. Lusque, L. Le Bescond, T. Filleron, Y. Pradat, A. Ducoulombier, B. Pistilli, T. Bachelot, F. Viret, C. Levy, N. Signolle, A. Alfaro, D. T. N. Tran, I. J. Garberis, H. Talbot, S. Christodoulidis, M. Vakalopoulou, N. Droin, V. Baris. Trastuzumab deruxtecan in metastatic breast cancer with variable HER2 expression: The phase 2 DAISY trial. Nature Medicine. Vol. 29, pg. 2110-2120, 2023, https://doi.org/10.1038/s41591-023-02478-2.

57 K. Witkowski, et al. Public perceptions of artificial intelligence in healthcare: A systematic review. Journal of Personalized Medicine. Vol. 13, pg. 110, 2023, https://doi.org/10.1186/s12910-024-01066-4.

58 E. Ratti, et al. Ethical and social considerations of applying artificial intelligence in healthcare. Bioethics. Vol. 37, pg. 15-25, 2023, https://doi.org/10.1111/bioe.13110.

59 S. Singh, et al. Socioeconomic impact of artificial intelligence on healthcare: A systematic review. Journal of Personalized Medicine. Vol. 13, pg. 110, 2023, https://doi.org/10.3390/jpm13010110.

60 M. Corpas, et al. Bridging genomics' greatest challenge: The diversity gap. PLOS Genetics. Vol. 18, pg. e1010200, 2022, https://doi.org/10.1371/journal.pgen.1010200.

© The National High School Journal of Science 2026