GMR
The Original

EMERGING LIQUID BIOPSY BIOMARKERS FOR EARLY CANCER DETECTION: SYSTEMATIC REVIEW OF CFDNA, CTDNA, METHYLATION, AND FRAGMENTOMICS

Faizan Akhtar[1]*, Shiva Krishna Rao T[2], Satish Tandale[3]

[1*]Assistant Professor, Department of Pathology, N.K.P. Salve Institute of Medical Sciences & Research Centre and Lata Mangeshkar Hospital, Nagpur, Maharashtra, India. Email: faizangmc3@gmail.com
[2] Senior Resident (2nd Year), Department of Clinical Pharmacology, Seth G.S. Medical College & KEM Hospital, Mumbai, Maharashtra, India.
[3] Assistant Professor, Department of Pathology, Government Medical College, Dharashiv, Maharashtra, India.

ABSTRACT
Background: Early cancer detection is central to reducing cancer-associated morbidity and mortality, yet established population screening is available for only a subset of malignancies. Liquid biopsy permits minimally invasive interrogation of tumor-derived molecular signals in blood, particularly circulating cell-free DNA (cfDNA) and its tumor-derived component, circulating tumor DNA (ctDNA).
Objective: To systematically review the biological basis, diagnostic performance, clinical evidence, strengths, limitations, and translational potential of cfDNA concentration/integrity, mutation-based ctDNA, DNA methylation, and fragmentomics for early and multi-cancer detection.
Methods: A structured literature search was undertaken for human studies published through June 2026. The search identified 676 records: 412 from PubMed/MEDLINE, 238 from other academic databases, and 26 through citation/reference searching. After removal of 124 duplicate records and 8 other records, 544 records underwent title and abstract screening; 454 were excluded. Ninety reports were sought for retrieval, 6 were not retrieved, and 84 full-text reports were assessed for eligibility. Sixty-six full-text reports were excluded for predefined reasons, leaving 18 primary studies in the qualitative synthesis. Reporting followed PRISMA 2020 and PRISMA-DTA principles, and diagnostic-accuracy studies were appraised using QUADAS-3. Owing to marked clinical and analytical heterogeneity, quantitative meta-analysis was not performed.
Results: Simple cfDNA concentration and integrity measures showed cancer-associated changes but limited tumor specificity. Mutation-based ctDNA assays offered high molecular specificity but were constrained by extremely low tumor fractions in stage I disease and by clonal hematopoiesis. Methylation-based assays provided a broader signal and strong cancer-signal-origin information; in an independent CCGA validation cohort, specificity was 99.5%, while sensitivity increased from 16.8% in stage I to 90.1% in stage IV disease. Fragmentomic approaches captured genome-wide differences in fragment size, coverage, nucleosome positioning, and end characteristics and demonstrated high discriminative performance with relatively shallow sequencing. Multimodal strategies generally improved information density. Among 17 diagnostic-accuracy studies evaluated with QUADAS-3, participant selection was the dominant concern: 13/17 (76.5%) had high risk in this domain. The 2026 NHS-Galleri randomized trial did not meet its prespecified primary endpoint for reducing combined stage III-IV cancers, although stage IV cancers decreased and stage I-II diagnoses increased.
Conclusion: Methylation and fragmentomics currently provide the most information-dense cfDNA-based approaches for early cancer detection, while mutation-based ctDNA retains high biological specificity. Multimodal integration is likely to provide the greatest diagnostic potential. However, high diagnostic accuracy does not establish population-screening benefit; randomized evidence demonstrating reductions in advanced cancer and mortality, with acceptable harms and costs, remains essential.

KEYWORDS: liquid biopsy; cell-free DNA; cfDNA; circulating tumor DNA; ctDNA; methylation; fragmentomics; multi-cancer early detection; cancer screening; precision oncology.

1. INTRODUCTION

Cancer survival is strongly influenced by stage at diagnosis. For many solid malignancies, localized disease can be treated with curative intent, whereas metastatic disease is associated with substantially poorer outcomes. Current population-based screening programs cover only a limited subset of cancers, leaving many high-mortality malignancies without an effective population-wide early-detection strategy. This gap has driven development of blood-based molecular tests capable of detecting multiple cancers from a single sample.
Liquid biopsy refers to the analysis of tumor-associated material circulating in body fluids, including circulating tumor cells, extracellular vesicles, RNA, proteins, metabolites, cfDNA, and ctDNA. Among these analytes, cfDNA is especially attractive because it can encode genetic, epigenetic, and structural information about both the presence and tissue origin of malignancy.

cfDNA consists predominantly of short DNA fragments released into the circulation during apoptosis, necrosis, and other biological processes. In healthy individuals, most plasma cfDNA is derived from hematopoietic cells. In cancer, a variable fraction originates from malignant cells and is termed ctDNA. This tumor-derived fraction can contain somatic mutations, copy-number abnormalities, aberrant methylation, and cancer-associated fragmentation signatures.

The principal challenge in early detection is biological rather than purely analytical: small localized tumors may contribute only a minute fraction of total plasma DNA. Bettegowda et al. demonstrated that ctDNA detectability varies considerably by cancer type and disease stage, with lower detection rates in localized disease [1]. Consequently, the field has moved from measuring only cfDNA quantity or searching for a limited set of mutations toward information-rich methylomic and fragmentomic approaches.

Cancer-associated methylation changes can occur across thousands of genomic loci and retain tissue-lineage information. Fragmentation is also non-random and reflects nucleosome positioning, chromatin accessibility, transcriptional activity, nuclease processing, and the tissue from which cfDNA originated. These properties have established methylomics and fragmentomics as major strategies for multi-cancer early detection (MCED).

The objective of this systematic review was to evaluate four major biomarker classes - cfDNA concentration/integrity, mutation-based ctDNA, DNA methylation, and fragmentomics - and to compare their biological rationale, diagnostic performance, clinical validation, risk of bias, and potential roles in population screening.

2. MATERIALS AND METHODS

2.1 Review design and reporting standard

This systematic review was prepared according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement and the PRISMA extension for Diagnostic Test Accuracy (PRISMA-DTA) [24,25]. Because the principal outcomes were diagnostic-accuracy measures, methodological quality was evaluated using QUADAS-3, the current version of the Quality Assessment of Diagnostic Accuracy Studies tool [26,27].

2.2 Information sources and search strategy

Structured searches were undertaken for literature published through June 2026 using PubMed/MEDLINE and multidisciplinary academic literature databases. Reference lists of major systematic reviews, translational reviews, and landmark validation studies were also examined. The principal search concept combined terms for liquid biopsy, early cancer detection, and the four biomarker domains.

Core search string: ("cell-free DNA" OR cfDNA OR "circulating tumor DNA" OR ctDNA) AND ("early cancer detection" OR "early-stage cancer" OR screening OR "multi-cancer early detection" OR MCED) AND (mutation OR sequencing OR methylation OR methylome OR fragmentomics OR fragmentation OR "fragment size" OR "nucleosome footprint" OR "end motif").

The evidence base was cross-checked against previous mapping and systematic reviews to reduce the likelihood of missing major primary studies [2,3,23].

2.3 Eligibility criteria

1. Human participants;
2. Plasma- or blood-derived cfDNA or ctDNA;
3. Cancer detection rather than exclusively prognosis, minimal residual disease, or treatment monitoring;
4. Early-stage/non-metastatic cancer, asymptomatic screening populations, prediagnostic samples, or appropriately defined non-cancer controls;
5. Assessment of cfDNA concentration/integrity, somatic mutations, methylation, fragmentomic features, or combinations of these biomarkers;
6. Reporting of at least one diagnostically relevant outcome, including sensitivity, specificity, AUC, PPV, negative predictive value, cancer-signal-origin accuracy, or prospective cancer-detection yield.

Animal studies, case reports, reviews without original patient data, studies restricted to therapeutic monitoring, and studies without interpretable diagnostic outcomes were excluded from the primary diagnostic evidence synthesis.

2.4 Study selection and data extraction

Titles and abstracts were evaluated for relevance, followed by full-text assessment of potentially eligible publications. The search identified 676 records, comprising 412 records from PubMed/MEDLINE, 238 from other academic databases, and 26 from citation/reference searching. Before screening, 124 duplicate records and 8 other records were removed, leaving 544 records for title and abstract screening. Of these, 454 records were excluded and 90 reports were sought for retrieval. Six reports could not be retrieved; therefore, 84 full-text reports were assessed for eligibility. Following full-text assessment, 18 primary studies were included in the qualitative synthesis. Data extracted included study design, population, cancer type, stage, biomarker technology, analytical method, sensitivity, specificity, AUC, PPV, cancer-signal-origin performance, and major limitations.

Among the 84 full-text reports assessed, 66 were excluded: 21 were not early-detection studies, 14 evaluated treatment monitoring or prognosis only, 9 investigated an inappropriate biomarker/analyte for the review question, 10 were reviews/editorials/commentaries without primary data, 7 did not provide sufficient diagnostic

outcomes, and 5 represented overlapping or duplicate patient populations. The final 18-study evidence set comprised 13/18 (72.2%) case-control, assay-development, or independent case-control validation studies; 4/18 (22.2%) prospective or longitudinal detection/screening cohort studies; and 1/18 (5.6%) randomized population-screening trial.

2.5 Risk-of-bias assessment

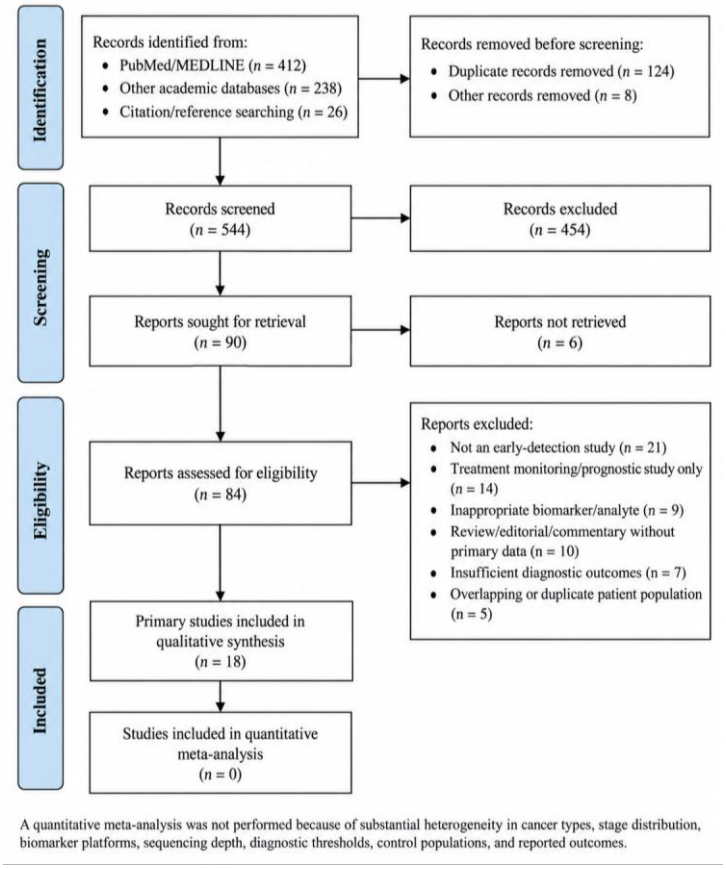
Diagnostic-accuracy studies were appraised using QUADAS-3 [26,27]. QUADAS-3 comprises six phases and assesses risk of bias across four domains: Participants, Index Test, Target Condition, and Analysis. The first three domains are also evaluated for concerns regarding applicability. Particular attention was given to case-control sampling, spectrum bias, prespecification of thresholds, separation of training and validation cohorts, independent external validation, clonal hematopoiesis filtering, reference standards, follow-up of test-negative participants, missing data, and model overfitting.

The 2026 NHS-Galleri study was interpreted separately as randomized clinical-utility evidence because its primary endpoint concerned stage distribution rather than conventional diagnostic-test accuracy. It was therefore not included in the numerical QUADAS-3 summary presented for the 17 diagnostic-accuracy studies.

2.6 Synthesis approach

A quantitative meta-analysis was not performed because of substantial heterogeneity in cancer type, disease stage, study design, sequencing depth, molecular platform, positivity threshold, control population, and reported outcomes. Results were therefore synthesized qualitatively and comparatively by biomarker class.

Figure 1. PRISMA 2020 flow diagram for study selection.



3. RESULTS

3.1 Study characteristics and PRISMA study selection

The database and supplementary searches identified 676 records. After 132 records were removed before screening (124 duplicates and 8 other records), 544 titles and abstracts were screened, of which 454 were excluded. Ninety reports were sought for retrieval; 6 were not retrieved, leaving 84 full-text reports for eligibility assessment. Sixty-six full-text reports were excluded for predefined reasons: not an early-detection study (n = 21), treatment monitoring/prognostic study only (n = 14), inappropriate biomarker/analyte (n = 9), review/editorial/commentary without primary data (n = 10), insufficient diagnostic outcomes (n = 7), and overlapping or duplicate patient population (n = 5). Consequently, 18 primary studies were included in the final qualitative synthesis, with no studies pooled in a quantitative meta-analysis because of substantial clinical and methodological heterogeneity. The included evidence covered cfDNA concentration/integrity, mutation-based ctDNA, methylation-based assays, fragmentomics, multimodal assays, prospective screening implementation, and randomized clinical utility. Thirteen studies used case-control or development/validation designs, four used prospective or longitudinal detection designs, and one evaluated population screening in a randomized trial.

Table 1. Characteristics of the 18 primary studies included in the principal evidence synthesis.

Study	Design / approach	Population	Key finding
Bettegowda et al., 2014 [1]	Mutation-based ctDNA	640 patients across multiple cancers	ctDNA detectability was strongly stage- and tumor-dependent; localized disease had lower detection rates.
Ren et al., 2023 [5]	cfDNA concentration + integrity + proteins	84 NSCLC; 50 controls	Combined model: sensitivity 94.05%, specificity 90.0%, AUC 0.968.
Phallen et al., 2017 [6]	TEC-Seq mutation detection	Early colorectal, breast, lung, ovarian cancers	Somatic alterations detectable in a substantial proportion of stage I-II cancers.
Cohen et al., 2018 [7]	CancerSEEK: mutations + proteins	1,005 cancers; 812 controls	Median sensitivity about 70%; specificity >99%; localization information provided.
Chabon et al., 2020 [8]	Integrated mutation features / Lung-CLiP	Early lung cancer and matched controls	Highlighted clonal hematopoiesis filtering and integrated feature analysis.
Shen et al., 2018 [9]	Plasma cfDNA methylomes	Multiple cancer types	Demonstrated sensitive cancer detection and classification using plasma methylation profiles.
Liu et al., 2020 [10]	Targeted methylation MCED	CCGA cohort	Demonstrated multicancer detection and cancer-signal localization.
Klein et al., 2021 [11]	Targeted methylation, independent validation	4,077 participants	Specificity 99.5%; sensitivity 16.8% stage I, 40.4% II, 77.0% III, 90.1% IV.
Chen et al., 2020 [12]	PanSeer methylation	Prediagnostic longitudinal samples	Cancer-associated methylation signal detected in samples obtained before conventional diagnosis.
Gao et al., 2023 [13]	THUNDER targeted methylation	1,010 independent validation participants	69.1% sensitivity, 98.9% specificity; 83.2% tissue-of-origin accuracy.
Jamshidi et al., 2022 [14]	Head-to-head cfDNA feature comparison	CCGA	Methylation and multimodal approaches were among the strongest for cancer detection/localization.
Mouliere et al., 2018 [15]	Fragment size analysis	344 plasma samples	Short-fragment selection enriched tumor-derived DNA and improved detection.
Cristiano et al., 2019 [16]	Genome-wide fragmentomics	236 cancers; 245 controls	Overall AUC 0.94; tissue-of-origin information obtained from fragmentation patterns.
Mathios et al., 2021 [17]	DELFI fragmentomics	Prospective high-risk lung cohort	Integrated approach detected 94% of cancers and 91% of stage I-II disease at 80% specificity.
Nguyen et al., 2023 [18]	SPOT-MAS methylomics + fragmentomics	738 nonmetastatic cancers; 1,550 controls	72.4% sensitivity at 97.0% specificity; multimodal shallow-sequencing strategy.
Lennon et al., 2020 [20]	Blood test + PET-CT	10,006 asymptomatic women	Prospective feasibility of blood-based multicancer detection with controlled diagnostic workup.
Schrag et al., 2023 [21]	PATHFINDER prospective MCED	6,621 evaluable participants	Cancer signal in 1.4%; PPV approximately 38%; median diagnostic resolution 79 days.
Swanton et al., 2026 [22]	NHS-Galleri randomized clinical-utility trial	142,924 adults aged 50-79 years	Primary stage III-IV endpoint not met; stage IV cancers decreased and stage I-II diagnoses increased.

3.2 Biological relationship between cfDNA and ctDNA

cfDNA represents the total extracellular DNA circulating in plasma, whereas ctDNA is the fraction released from malignant cells. Tumor-derived DNA can differ from background cfDNA in sequence, copy number, methylation state, fragment size, nucleosomal organization, preferred fragment endpoints, and other structural characteristics.

The ctDNA fraction is influenced by tumor burden, vascularity, anatomic site, biological aggressiveness, cell turnover, and metastatic spread. Consequently, assays that perform well in advanced disease often show lower sensitivity in stage I cancer. In low-prevalence screening populations, specificity becomes equally important because even a modest false-positive rate can generate many unnecessary diagnostic investigations.

3.3 cfDNA concentration and DNA integrity

One of the earliest liquid-biopsy approaches was measurement of total cfDNA concentration. Cancer can increase circulating DNA through apoptosis, necrosis, altered clearance, and increased cellular turnover. DNA integrity indices compare relatively long and short fragments, often using repetitive ALU sequences.

A systematic review of cfDNA integrity found substantial heterogeneity, including inconsistent direction of effect between studies [4]. Ren et al. evaluated 84 patients with non-small-cell lung cancer and 50 healthy controls; a model combining cfDNA concentration, integrity, CEA, CA125, NSE, and CYFRA21-1 achieved 94.05% sensitivity, 90.0% specificity, and an AUC of 0.968 [5]. However, the signal was stronger in more advanced disease.

Total cfDNA is also elevated in non-malignant conditions including inflammation, infection, trauma, ischemia, strenuous exercise, autoimmune disease, and pregnancy. Accordingly, cfDNA concentration and simple integrity ratios lack sufficient biological specificity for stand-alone population screening and are better considered supplementary features within multimodal classifiers.

3.4 Mutation-based ctDNA detection

Mutation-based assays offer high biological specificity because detection of a validated somatic driver alteration can indicate a clonal neoplastic process. Digital PCR, error-corrected next-generation sequencing, CAPP-Seq, and related approaches have progressively lowered detectable variant allele fractions.

Bettegowda et al. showed that ctDNA is detectable across many malignancies but that sensitivity is lower in localized disease [1]. Phallen et al. subsequently used targeted error-corrected sequencing and demonstrated that somatic alterations could be identified in a substantial proportion of stage I-II colorectal, breast, lung, and ovarian cancers [6].

CancerSEEK combined mutations with circulating proteins. In 1,005 patients with eight non-metastatic cancers and 812 controls, median sensitivity was approximately 70% and specificity exceeded 99% [7]. Chabon et al. demonstrated that integration of genomic features improves early lung-cancer detection and emphasized the need to distinguish tumor mutations from clonal hematopoiesis [8].

Clonal hematopoiesis is particularly important in older screening populations. Mutations in genes such as DNMT3A, TET2, ASXL1, and TP53 can arise from expanded hematopoietic clones and appear in plasma without an occult solid tumor. Paired leukocyte sequencing, mutation-context analysis, and integration of orthogonal signals can reduce this source of false-positive classification.

Another limitation is sampling. A small localized tumor may release so few mutant molecules that a standard blood draw contains only a very small number, or none, of the particular mutant copies being targeted. Increasing sequencing depth cannot recover molecules absent from the sampled plasma. This biological ceiling has encouraged methods that interrogate thousands to millions of genomic features.

3.5 cfDNA methylation biomarkers

DNA methylation is among the most promising sources of signal for early cancer detection. Aberrant hypermethylation and hypomethylation emerge during carcinogenesis, can involve large numbers of genomic loci, and frequently preserve tissue-specific patterns.

Shen et al. demonstrated that plasma cfDNA methylomes generated using an enrichment-based method could discriminate multiple cancer types and classify their origins [9]. The CCGA program subsequently provided large-scale validation. Liu et al. showed that targeted methylation signatures could detect numerous cancers and predict cancer-signal origin [10].

In an independent validation cohort of 4,077 participants, Klein et al. reported 99.5% specificity and overall sensitivity of 51.5%. Sensitivity was strongly stage dependent: 16.8% for stage I, 40.4% for stage II, 77.0% for stage III, and 90.1% for stage IV disease. Among true-positive results, cancer-signal origin was correctly predicted in 88.7% [11].

PanSeer provided evidence that methylation signals may be detectable before conventional diagnosis. Chen et al. identified cancer-associated signatures in stored plasma obtained before diagnosis, although the authors emphasized the need for prospective confirmation [12]. THUNDER used a targeted methylation panel and, in an independent validation cohort, achieved 69.1% sensitivity and 98.9% specificity with 83.2% tissue-of-origin accuracy [13].

Direct comparison of cfDNA features in CCGA indicated that methylation is highly informative for both cancer-signal detection and localization [14]. The principal limitation remains stage I sensitivity, demonstrating that even information-rich methylation assays are constrained by extremely low tumor-derived signal in the earliest cancers.

3.6 Fragmentomics

Fragmentomics analyzes structural properties of circulating DNA rather than focusing solely on nucleotide sequence. Features include fragment-length distributions, short-to-long ratios, genome-wide coverage patterns, preferred endpoints, end motifs, nucleosome footprints, orientation-aware fragmentation, promoter fragmentation patterns, and copy-number-linked fragmentation.

Tumor-derived cfDNA is enriched within particular size ranges and exhibits altered fragmentation compared with hematopoietic cfDNA. Mouliere et al. demonstrated enrichment of ctDNA among shorter fragments and showed that size selection could increase the relative representation of tumor-derived DNA [15].

Cristiano et al. developed a genome-wide fragmentation approach in 236 patients with seven cancer types and 245 healthy individuals. Their classifier achieved an overall AUC of 0.94, and fragmentation patterns also provided tissue-of-origin information [16]. Mathios et al. prospectively evaluated the DELFI approach in individuals at risk for lung cancer; an integrated model detected 94% of cancers overall, including 91% of stage I-II cancers, at 80% specificity [17].

Fragmentomics has important conceptual advantages: it surveys signals distributed across much of the genome, can operate with relatively shallow sequencing, and does not require a recurrent driver mutation to be present in the sampled molecules. Its principal challenges are bioinformatic complexity, batch effects, model transportability, pre-analytical standardization, and the need for robust external validation.

3.7 Multimodal liquid biopsy

No individual cfDNA feature fully captures the biological diversity of cancer. Accordingly, current development increasingly combines multiple molecular dimensions. SPOT-MAS integrates methylation, fragment size, copy-number changes, and fragment-end motifs. In a cohort of 738 nonmetastatic cancers and 1,550 controls, it achieved 72.4% sensitivity at 97.0% specificity [18].

CancerSEEK provides another multimodal paradigm by integrating somatic mutations with circulating proteins [7]. Other strategies combine mutations, methylation, copy-number abnormalities, fragmentomic features, conventional biomarkers, and clinical risk variables.

A 2026 systematic review and meta-analysis of 109 studies integrating machine learning with cfDNA found consistently high specificity and reported pooled sensitivities of approximately 86% for fragmentation-based classifiers and 81% for methylation-based classifiers, although marked heterogeneity existed across tumor types, stages, algorithms, and populations [19]. These results support multimodal integration while reinforcing the need for prospective validation.

Table 2. Comparative characteristics of major cfDNA-based biomarker strategies.

Biomarker	Biological signal	Major advantages	Principal limitations	Likely role
Total cfDNA concentration	Increased DNA release into plasma	Simple; inexpensive; rapid	Poor cancer specificity; affected by inflammation, injury, ischemia and other conditions	Supplementary feature
cfDNA integrity	Altered long/short fragment ratio	Technically straightforward	Conflicting results; assay and pre-analytical heterogeneity	Component of multivariable models
Mutation-based ctDNA	Tumor-associated somatic variants	High molecular specificity; genomic actionability	Low tumor fraction in early disease; clonal hematopoiesis; limited recurrent targets	Targeted detection and multimodal assays
DNA methylation	Cancer- and tissue-specific CpG alterations	Many informative loci; strong tissue-of-origin signal	Assay complexity; biological confounding; stage-dependent sensitivity	Major MCED platform
Fragmentomics	Fragment length, coverage, ends, motifs and nucleosomal architecture	Genome-wide; relatively shallow sequencing; mutation independent	Bioinformatic complexity; batch effects; standardization	MCED and multi-omic integration
Multimodal assays	Combined genomic, epigenomic, structural and/or protein signals	Higher information density; complementary biology	Cost; computational complexity; overfitting and validation burden	Next-generation screening strategy

3.8 Prospective screening evidence

Case-control diagnostic performance does not necessarily predict screening performance. The intended population is predominantly cancer-free and includes many individuals with benign disease, age-related clonal changes, and chronic comorbidity that may be underrepresented in healthy-control datasets.

The DETECT-A study enrolled 10,006 women without a prior cancer diagnosis and demonstrated prospective feasibility of combining a blood test with PET-CT-guided diagnostic evaluation. Twenty-six cancers were detected by blood testing, while only a small proportion of participants underwent PET-CT because of false-positive blood tests [20].

PATHFINDER enrolled 6,662 adults aged 50 years or older. Among 6,621 evaluable participants, a cancer signal was detected in 92 (1.4%); 35 subsequently had cancer diagnosed, yielding a positive predictive value of

approximately 38%. Median time to diagnostic resolution was 79 days, illustrating the importance of downstream diagnostic burden and time-to-resolution metrics [21].

A 2025 systematic review of general-population MCED screening found promising specificity but substantial variation in sensitivity and limited evidence on mortality, quality of life, overdiagnosis, harms, and cost-effectiveness [2].

3.9 NHS-Galleri randomized screening trial

The NHS-Galleri randomized trial enrolled 142,924 asymptomatic adults aged 50-79 years in England and evaluated annual methylation-based MCED testing over three screening rounds [22]. The prespecified primary endpoint - a statistically significant reduction in combined stage III-IV incidence across 12 prespecified cancers - was not met (706 versus 688 cases; incidence-rate ratio 1.03). However, stage IV cancers were reduced by 14% (342 versus 397), stage I-II diagnoses increased by 16%, and screen-detected cancers increased substantially. Aggregate test specificity was 99.55% and positive predictive value was 52.0% [22].

These findings are central to interpretation of MCED evidence because they demonstrate that high diagnostic performance is not equivalent to proven clinical benefit. Stage shift, mortality, overdiagnosis, diagnostic burden, and adverse consequences must be evaluated directly in randomized screening programs.

3.10 QUADAS-3 risk-of-bias assessment

Seventeen diagnostic-accuracy studies were assessed using QUADAS-3. Participant selection was the dominant methodological concern. Thirteen of 17 studies (76.5%) were judged at high risk in the Participants domain because they relied on clinically diagnosed cancer cases compared with cancer-free or healthy controls, creating potential spectrum bias relative to the intended asymptomatic screening population. Three studies had some concerns/uncertainty and one had low risk in this domain.

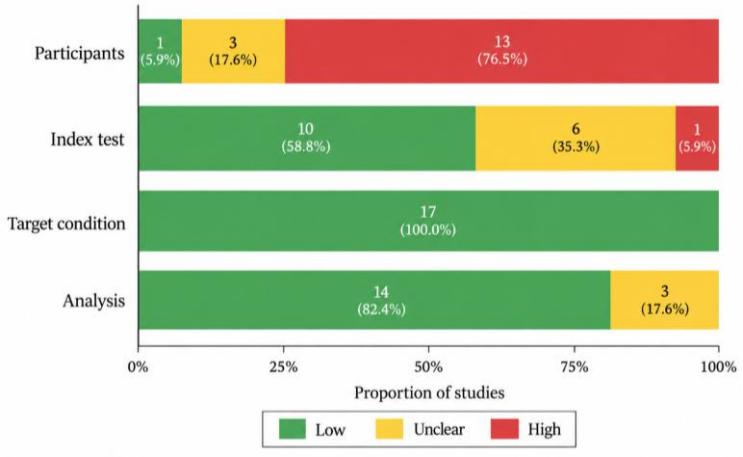
The Index Test domain was judged low risk in 10/17 studies (58.8%), some concerns/unclear in 6/17 (35.3%), and high risk in 1/17 (5.9%). The Target Condition domain was generally low risk because cancer status was usually established through histopathology, imaging, registries, clinical investigation, or combinations of these methods. The Analysis domain was low risk in 14/17 studies (82.4%) and had some concerns in 3/17 (17.6%), principally because of incomplete follow-up, machine-learning optimization, or uncertainty regarding handling of indeterminate and missing data.

Table 3. QUADAS-3 risk-of-bias assessment of the 17 diagnostic-accuracy studies.

Study	Participants	Index test	Target condition	Analysis	Overall interpretation
Bettegowda et al., 2014 [1]	High	Some concerns	Low	Low	High
Ren et al., 2023 [5]	High	High	Low	Some concerns	High
Phallen et al., 2017 [6]	High	Some concerns	Low	Low	High
Cohen et al., 2018 [7]	High	Some concerns	Low	Low	High
Chabon et al., 2020 [8]	High	Low	Low	Low	High
Shen et al., 2018 [9]	High	Some concerns	Low	Low	High
Liu et al., 2020 [10]	High	Low	Low	Low	High
Klein et al., 2021 [11]	High	Low	Low	Low	High
Chen et al., 2020 [12]	Some concerns	Low	Low	Low	Some concerns
Gao et al., 2023 [13]	High	Low	Low	Low	High
Jamshidi et al., 2022 [14]	High	Low	Low	Low	High
Mouliere et al., 2018 [15]	High	Some concerns	Low	Low	High
Cristiano et al., 2019 [16]	High	Some concerns	Low	Low	High
Mathios et al., 2021 [17]	Some concerns	Low	Low	Low	Some concerns
Nguyen et al., 2023 [18]	High	Low	Low	Low	High
Lennon et al., 2020 [20]	Some concerns	Low	Low	Some concerns	Some concerns
Schrag et al., 2023 [21]	Low	Low	Low	Some concerns	Some concerns

Note: QUADAS-3 is designed for estimate-level appraisal. The table summarizes the principal accuracy estimate(s) contributing to this review and should not be interpreted as a numerical quality score. The NHS-Galleri randomized clinical-utility trial [22] was not included in this DTA summary.

Figure 2. QUADAS-3 risk-of-bias summary across the 17 diagnostic-accuracy studies.



4. DISCUSSION

4.1 Principal findings

The evolution of cfDNA-based cancer detection can be understood as a transition from measuring the quantity of circulating DNA to decoding its biological information. Total cfDNA concentration and simple integrity indices provided early proof that circulating DNA differs in patients with cancer but lack sufficient specificity. Mutation-based ctDNA substantially increases molecular specificity but exposes the fundamental biological limit of low tumor-derived DNA abundance in the earliest cancers.

Methylation and fragmentomics partially address this limitation by expanding the number of informative features. Methylation provides strong tissue-lineage information and can interrogate thousands of differentially methylated regions. Fragmentomics captures genome-wide consequences of altered chromatin architecture and nuclease processing. These approaches therefore extract more information from each available DNA molecule than simple mutation counting.

The current evidence increasingly supports multimodal analysis. However, the risk-of-bias assessment demonstrates that most diagnostic-development studies are not direct simulations of population screening. High performance in enriched case-control cohorts should therefore be interpreted as evidence of analytical and biological potential rather than proof of population-level effectiveness.

4.2 The stage I sensitivity problem

The most difficult biological challenge remains stage I cancer. In the large CCGA validation study, targeted methylation achieved only 16.8% sensitivity for stage I disease despite 99.5% specificity [11]. This illustrates the paradox of early detection: the cancers for which screening could provide the greatest therapeutic advantage are often those shedding the least tumor-derived material. Improvement may require larger plasma inputs, optimized molecular recovery, serial longitudinal testing, orthogonal biomarker integration, and individualized risk models rather than sequencing depth alone.

4.3 Specificity, prevalence, and positive predictive value

For population screening, extremely high specificity is essential. When cancer prevalence is low, even a small false-positive percentage can generate many unnecessary scans, biopsies, and episodes of anxiety. PATHFINDER illustrates this distinction: despite high analytical specificity, only 35 of 92 participants with a cancer signal were ultimately diagnosed with cancer [21]. Screening performance must therefore be evaluated using PPV, diagnostic burden, time to resolution, procedure-related harms, and cost alongside sensitivity and specificity.

4.4 Tissue-of-origin prediction

A practical MCED test must identify not only whether a cancer signal is present but also where it is likely to originate. Methylation has a particular advantage because tissue-specific epigenetic programming is partly retained in tumor-derived DNA. CCGA reported 88.7% cancer-signal-origin accuracy among true positives, while THUNDER reported 83.2% tissue-of-origin accuracy [11,13]. Fragmentation and nucleosomal footprints can provide complementary tissue information because chromatin organization reflects lineage and transcriptional state.

4.5 Clonal hematopoiesis and biological false positives

Mutation-based detection is particularly vulnerable to clonal hematopoiesis. Age-related expansion of hematopoietic clones can generate cancer-associated mutations in plasma even when no solid tumor is present. Paired leukocyte sequencing, fragment characteristics, methylation context, and tissue-specific signals can help distinguish hematopoietic from solid-tumor origins. More broadly, an analytically real molecular abnormality is

not necessarily evidence of a clinically significant cancer; assays must distinguish malignant signals from benign clonal, inflammatory, and premalignant processes.

4.6 Overdiagnosis, lead-time bias, and clinical utility

Earlier diagnosis does not automatically reduce mortality. Screening can identify indolent tumors that would never have caused symptoms, leading to overdiagnosis and unnecessary treatment. Lead-time bias can also make survival after diagnosis appear longer without delaying death. Therefore, stage shift is an intermediate endpoint rather than definitive proof of benefit. The NHS-Galleri findings reinforce this principle: the primary combined stage III-IV endpoint was not met even though several secondary findings were favorable [22].

4.7 Artificial intelligence and machine learning

Modern methylomic and fragmentomic assays generate high-dimensional data that require machine-learning methods. Neural networks, random forests, and ensemble approaches can integrate many weak signals. However, they also introduce risks of overfitting, data leakage, batch-specific signatures, demographic confounding, poor calibration, and limited transportability. Training, validation, and final test sets must therefore be strictly separated, and prospective external validation is essential. The 2026 meta-analysis by Filis et al. reported high diagnostic performance but considerable heterogeneity across tumor types, stages, platforms, and modelling approaches [19].

4.8 Implications of risk of bias

The QUADAS-3 findings materially affect interpretation of the literature. The predominance of high risk in the Participants domain reflects the widespread use of clinically apparent cancers and healthy controls in early assay development. These designs are efficient for discovery but can exaggerate apparent discrimination compared with unselected asymptomatic populations. Later prospective studies improve external relevance by reporting PPV, diagnostic resolution, and downstream investigations. Randomized screening trials provide an even higher level of clinical-utility evidence because they test whether molecular detection translates into meaningful stage shifts and, ultimately, improved patient outcomes.

5. Clinical and Translational Challenges

Pre-analytical standardization: Blood collection tube type, processing interval, centrifugation, plasma storage, freeze-thaw cycles, extraction method, and cfDNA input influence analytical results.

Analytical sensitivity: Early-stage tumors may produce tumor fractions near or below the limits of reliable molecular detection.

Biological heterogeneity: ctDNA shedding differs markedly across tumor types, anatomical sites, tumor volumes, and biological phenotypes.

Population calibration: Models trained using clinically diagnosed cases and exceptionally healthy controls may perform less well in older screening populations with chronic disease.

Diagnostic pathways: A positive MCED result requires a standardized, rapid, evidence-based pathway for localization, imaging, and tissue confirmation.

Cost-effectiveness: Sequencing, computational analysis, repeat testing, imaging, biopsies, and false-positive investigations must be included in economic evaluation.

Health equity: Validation should include diverse ancestry, geography, socioeconomic status, age, sex, and comorbidity profiles.

Clinical utility: Tests must demonstrate a favorable balance among lives saved, morbidity avoided, false-positive procedures, overdiagnosis, cost, and psychological consequences.

6. Future Directions

The next generation of liquid-biopsy screening is likely to be multidimensional rather than single-analyte. Integrated methylation-fragmentation models are biologically complementary because methylation captures epigenetic reprogramming while fragmentation reflects chromatin architecture and DNA-processing biology.

Longitudinal within-person monitoring may increase sensitivity by identifying molecular deviation from an individual baseline rather than relying exclusively on cross-sectional comparison with a population reference. Individualized models incorporating age, smoking, inherited susceptibility, family history, imaging findings, and previous screening results may further improve predictive value.

Technical development should prioritize efficient molecular recovery and robust low-input sequencing rather than simply increasing sequencing depth. Algorithms should be prospectively locked, externally validated, calibrated across populations, and reported transparently.

Most importantly, research priorities should shift from retrospective diagnostic-accuracy studies toward randomized screening trials. Critical endpoints include stage IV incidence, cancer-specific mortality, all-cause mortality where feasible, quality of life, diagnostic complications, overdiagnosis, healthcare utilization, and cost-effectiveness.

7. Limitations of the Review and Current Evidence

The literature is clinically and analytically heterogeneous. Cancer types, disease stages, sample-processing methods, sequencing depth, machine-learning algorithms, and positivity thresholds vary substantially, limiting direct numerical comparisons across studies.

Many studies used enriched case-control designs and therefore may overestimate performance relative to population screening. Relatively few studies have evaluated asymptomatic individuals prospectively, and randomized evidence remains limited.

The present review did not perform a pooled meta-analysis because the underlying studies were not sufficiently homogeneous in population, index test, threshold, cancer spectrum, stage distribution, control selection, and outcome definition.

8. CONCLUSION

Liquid biopsy has shifted cancer detection from organ-specific screening toward molecular identification of multiple malignancies using a single blood sample. Simple cfDNA concentration and integrity measures are insufficiently specific for stand-alone screening. Mutation-based ctDNA provides highly specific genomic evidence but is limited by low tumor-derived DNA abundance and clonal hematopoiesis. Methylation profiling expands the available signal and offers strong tissue-of-origin information, while fragmentomics captures genome-wide structural signatures without requiring predefined tumor mutations.

The most promising future strategy is likely to integrate complementary genomic, epigenomic, and fragmentomic information using rigorously validated computational models. Nevertheless, diagnostic accuracy is not synonymous with screening benefit. The NHS-Galleri trial demonstrates that favorable molecular performance and secondary stage-shift findings can coexist with failure to meet a prespecified primary clinical endpoint. Widespread implementation should therefore depend on prospective evidence of meaningful reductions in advanced cancer and mortality, together with acceptable harms, diagnostic burden, cost, and equitable performance.

DECLARATIONS

Funding: No specific funding was received for preparation of this review.

Conflict of Interest: The authors declare no conflict of interest.

Ethical Approval: Not applicable because this review used previously published data and did not involve direct participation of human subjects.

Informed Consent: Not applicable.

Data Availability: All information synthesized in this review was obtained from published literature.

REFERENCES

1. Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med. 2014;6(224):224ra24. doi:10.1126/scitranslmed.3007094.

2. Wade R, Nevitt S, Liu Y, Harden M, Khouja C, Raine G, et al. Multi-cancer early detection tests for general population screening: a systematic literature review. Health Technol Assess. 2025;29(2):1-105. doi:10.3310/DLMT1294.

3. Rubinstein WS, Patriotis C, Dickherber A, et al. Cancer screening with multicancer detection tests: a translational science review. CA Cancer J Clin. 2024;74(4):368-382. doi:10.3322/caac.21833.

4. Cheng J, Tang Q, Cao X, et al. Cell-free circulating DNA integrity based on peripheral blood as a biomarker for diagnosis of cancer: a systematic review. Cancer Epidemiol Biomarkers Prev. 2017;26(11):1595-1602. doi:10.1158/1055-9965.EPI-17-0502.

5. Ren S, Zeng G, Yi Y, et al. Combinations of plasma cfDNA concentration, integrity and tumor markers are promising biomarkers for early diagnosis of non-small cell lung cancer. Heliyon. 2023;9(10):e20851. doi:10.1016/j.heliyon.2023.e20851.

6. Phallen J, Sausen M, Adleff V, et al. Direct detection of early-stage cancers using circulating tumor DNA. Sci Transl Med. 2017;9(403):eaan2415. doi:10.1126/scitranslmed.aan2415.

7. Cohen JD, Li L, Wang Y, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science. 2018;359(6378):926-930. doi:10.1126/science.aar3247.

8. Chabon JJ, Hamilton EG, Kurtz DM, et al. Integrating genomic features for non-invasive early lung cancer detection. Nature. 2020;580(7802):245-251. doi:10.1038/s41586-020-2140-0.

9. Shen SY, Singhania R, Fehringer G, et al. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. Nature. 2018;563(7732):579-583. doi:10.1038/s41586-018-0703-0.

10. Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV; CCGA Consortium. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Ann Oncol. 2020;31(6):745-759. doi:10.1016/j.annonc.2020.02.011.

11. 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. Ann Oncol. 2021;32(9):1167-1177. doi:10.1016/j.annonc.2021.05.806.

12. Chen X, Gole J, Gore A, et al. Non-invasive early detection of cancer four years before conventional diagnosis using a blood test. Nat Commun. 2020;11:3475. doi:10.1038/s41467-020-17316-z.

13. Gao Q, Lin YP, Li BS, et al. Unintrusive multi-cancer detection by circulating cell-free DNA methylation sequencing (THUNDER): development and independent validation studies. Ann Oncol. 2023;34(5):486-495. doi:10.1016/j.annonc.2023.02.010.

14. Jamshidi A, Liu MC, Klein EA, et al. Evaluation of cell-free DNA approaches for multi-cancer early detection. Cancer Cell. 2022;40(12):1537-1549.e12. doi:10.1016/j.ccell.2022.10.022.

15. Mouliere F, Chandrananda D, Piskorz AM, et al. Enhanced detection of circulating tumor DNA by fragment size analysis. Sci Transl Med. 2018;10(466):eaat4921. doi:10.1126/scitranslmed.aat4921.
16. Cristiano S, Leal A, Phallen J, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. Nature. 2019;570(7761):385-389. doi:10.1038/s41586-019-1272-6.
17. Mathios D, Johansen JS, Cristiano S, et al. Detection and characterization of lung cancer using cell-free DNA fragmentomes. Nat Commun. 2021;12:5060. doi:10.1038/s41467-021-24994-w.
18. Nguyen VTC, Nguyen TH, Doan NNT, et al. Multimodal analysis of methylomics and fragmentomics in plasma cell-free DNA for multi-cancer early detection and localization. eLife. 2023;12:RP89083. doi:10.7554/eLife.89083.
19. Filis P, Markozannes G, Salgkamis D, et al. Integrating liquid biopsies and artificial intelligence for early cancer detection: a systematic review and meta-analysis. Eur J Cancer. 2026;239:116699. doi:10.1016/j.ejca.2026.116699.
20. Lennon AM, Buchanan AH, Kinde I, et al. Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. Science. 2020;369(6499):eabb9601. doi:10.1126/science.abb9601.
21. Schrag D, Beer TM, McDonnell CH III, et al. Blood-based tests for multicancer early detection (PATHFINDER): a prospective cohort study. Lancet. 2023;402(10409):1251-1260. doi:10.1016/S0140-6736(23)01700-2.
22. Swanton RC, Johnson P, Round T, et al. NHS-Galleri: Primary results from a randomised controlled trial to assess the clinical utility of a multi-cancer early detection (MCED) test in population screening. J Clin Oncol. 2026;44(17_suppl):LBA100. doi:10.1200/JCO.2026.44.17_suppl.LBA100.
23. Cree IA, Uttley L, Woods HB, et al. The evidence base for circulating tumour DNA blood-based biomarkers for the early detection of cancer: a systematic mapping review. BMC Cancer. 2017;17:697. doi:10.1186/s12885-017-3693-7.
24. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi:10.1136/bmj.n71.
25. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM; PRISMA-DTA Group. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: the PRISMA-DTA statement. JAMA. 2018;319(4):388-396. doi:10.1001/jama.2017.19163.
26. Whiting PF, Tomlinson E, Rutjes AWS, et al. QUADAS-3: a revised tool for the quality assessment of diagnostic test accuracy studies. Ann Intern Med. 2026;179(4):548-555. doi:10.7326/ANNALS-25-02104.
27. Davenport CF, Rutjes AWS, Mallett S, et al. QUADAS-3 Explanation and Elaboration: guidance for quality assessment of diagnostic test accuracy studies. Ann Intern Med. 2026;179(4):e2504943. doi:10.7326/ANNALS-25-04943.