

Title: Clinical validity and utility of circulating tumor DNA (ctDNA) testing in advanced non-small cell lung cancer (aNSCLC): a systematic literature review and meta-analysis

Authors (in order): Cheng Chen, PhD¹; Michael P. Douglas, MS¹; Meera V. Ragavan, MD, MPH²;

Kathryn A. Phillips, PhD^{1,3,4}; and Jeroen P. Jansen, PhD^{1,3,4}

Author Affiliations:

1. Department of Clinical Pharmacy, UCSF Center for Translational and Policy Research on Precision Medicine (TRANSPERS), San Francisco, California

2. Division of Hematology and Oncology, UCSF Department of Medicine, San Francisco, California

3. UCSF Helen Diller Family Comprehensive Cancer Center, San Francisco, California

4. UCSF Philip R. Lee Institute for Health Policy, San Francisco, California

Corresponding Author: Jeroen Jansen, Department of Clinical Pharmacy, School of Pharmacy, University of California San Francisco, 490 Illinois St. Valley Tower, 3rd Floor, Box 0613, San Francisco, CA 94143 (jeroen.jansen@ucsf.edu)

Funding: This work was supported by grants from the National Cancer Institute (R01 CA221870) and the National Human Genome Research Institute (U01 HG009599 and HG011792) to Dr. Phillips.

Disclaimer: The National Human Genome Research Institute and the National Cancer Institute had no role in the preparation, review, or approval of the manuscript or decision to submit the manuscript for publication.

ABSTRACT

Purpose: Circulating tumor DNA (ctDNA) testing has become a promising tool to guide first-line (1L) targeted treatment for advanced non-small cell lung cancer (aNSCLC). This study aims to estimate the clinical validity (CV) and clinical utility (CU) of ctDNA-based next-generation sequencing (NGS) for oncogenic driver mutations to inform 1L treatment decisions in aNSCLC through a systematic literature review and meta-analysis.

Methods: A systematic literature search was conducted in PubMed/MEDLINE and Embase to identify randomized control trials or observational studies reporting CV/CU on ctDNA testing in patients with aNSCLC. Meta-analyses were performed using bivariate random-effects models to estimate pooled sensitivity and specificity. Progression-free/overall survival (PFS/OS) was summarized for CU studies.

Results: Eighteen studies were identified: 17 CV only, 2 CU only, and 1 both. Thirteen studies were included for the meta-analysis on multi-gene detection. The overall sensitivity and specificity for ctDNA detection of any mutation were 0.69 (95% CI, 0.63–0.74) and 0.99 (95% CI, 0.97–1.00) respectively. However, sensitivity varied greatly by driver gene, ranging from 0.29 (95% CI, 0.13–0.53) for *ROS1* to 0.77 (95% CI, 0.63–0.86) for *KRAS*. Two studies compared PFS with ctDNA versus tissue-based testing followed by 1L targeted therapy found no significant differences. One study reported OS curves on ctDNA-matched and tissue-matched therapies but no hazard ratios were provided.

Conclusion: ctDNA testing demonstrated an overall acceptable diagnostic accuracy in aNSCLC patients, however, sensitivity varied greatly by driver mutation. Further research is needed, especially for uncommon driver mutations, to better understand the CU of ctDNA testing in guiding targeted treatments for aNSCLC.

Keywords: ctDNA, NSCLC, sensitivity, specificity, meta-analysis

Introduction

Non-small cell lung cancer (NSCLC) accounts for an estimated 85% of lung cancer cases, with adenocarcinoma, squamous cell carcinoma, and large cell carcinoma being the most common histological subtypes.¹ The estimated five-year survival is only about 6% in patients with advanced-stage disease.^{2,3} In the past decade, more than 20 targeted therapies have been approved for the treatment of advanced NSCLC (aNSCLC) in patients who harbor *EGFR*, *BRAF*, *MET*, *RET*, *NTRK*, *KRAS*, *ALK*, or *ROS1* alterations.⁴ Targeted treatment according to the presence of oncogenic driver mutations has been associated with improved survival outcomes⁴⁻⁶, and current National Comprehensive Cancer Network (NCCN) guidelines specifically recommend that all patients with aNSCLC should get broad genomic profiling with next-generation sequencing (NGS), given its more optimal use of sample availability, reduced procedure time, and favorable testing costs compared to single gene tests.⁷⁻⁹

In recent years, there has been an increase in the use of circulating tumor DNA (ctDNA) from blood samples as an alternative to tissue biopsy (TB) for identifying oncogenic driver mutations to inform first-line (1L) aNSCLC therapy.¹⁰ Advantages of ctDNA testing include the avoidance of an invasive procedure and possible complications, the ability to identify driver mutations for patients with limited tissue availability for comprehensive NGS, and shorter turnaround time allowing for faster initiation of 1L therapy.¹¹ Several ctDNA-based NGS tests have been developed and approved for NSCLC, including Guardant360[®] CDx and FoundationOne[®] Liquid CDx.^{12,13}

However, the diagnostic accuracy of ctDNA testing remains unclear due to variations in technologies and use scenarios.^{14,15} With the broadened adoption of ctDNA-based NGS tests in NSCLC, a better understanding of their clinical validity (CV) and clinical utility (CU) is needed. “Clinical validity” usually describes the diagnostic accuracy of genetic tests and is generally measured by sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).^{16,17} Without proper validation of test performance, accurate identification of mutations may be compromised, leading to delays in suitable

clinical decisions. Previous systematic literature reviews have examined the CV of ctDNA testing in aNSCLC patients but findings have been inconsistent due to varied review focus and inclusion criteria.¹⁸⁻

²⁶ Most of the previous systematic literature reviews and/or meta-analyses^{18,19,22,23,25,26} did not specify sequencing technologies, which could potentially explain the observed heterogeneity in findings and limit their clinical implications in relation to the currently more recommended NGS tests.

“Clinical utility” refers to the risks or benefits from test use and in our study is measured as progression-free or overall survival (PFS/OS), considering the therapeutic effectiveness outcomes are the most relevant and commonly reported measure in CU studies and that the focus of our review is 1L targeted treatment informed by ctDNA testing.²⁷⁻²⁹ CU evidence directly answers if a test is useful in improving patient health outcomes – it is an essential element in value evaluation and affects the acceptance of a test from all parties (patients, healthcare providers and payers). To our knowledge, no study has systematically reviewed the CU of ctDNA testing for a comprehensive set of biomarkers to inform 1L treatment decisions in aNSCLC.

Given the rapidly evolving field of ctDNA testing technologies and the limitations of previous review studies, an updated synthesis of CV and CU evidence is warranted. The objective of the current study was to estimate the CV and CU of ctDNA-based NGS for oncogenic driver mutations to inform 1L treatment decisions in aNSCLC by means of a systematic literature review and meta-analysis of currently available evidence.

Materials and Methods

Systematic literature review

Eligibility criteria

The study inclusion criteria were defined in terms of the population, interventions, comparisons, outcomes, and study design (PICOS) to guide the identification and selection of relevant studies:

Population - Adult patients with aNSCLC (at least 80% with stage III or IV thereby making the assumption that reported results based on the total study population are still applicable to the stage III or IV NSCLC target population of interest), with a subset of treatment-naïve patients for CV studies and all being treatment-naïve for CU studies; *Interventions* – ctDNA-based NGS especially for the detection of clinically relevant driver mutations (i.e., *EGFR*, *BRAF*, *MET*, *RET*, *NTRK*, *KRAS*, *ALK*, *ROS1*), and when outcome of interest is CU at least 80% of population receiving matched targeted treatment by driver mutation according to current NCCN guidelines⁷; *Comparators for CU studies* – ctDNA testing vs. TB; *Outcomes in CV studies* – sensitivity, specificity, PPV, NPV or any other measures that allow calculation of true positive (TP), false positive (FP), false negative (FN), and true negative (TN) rates; *Outcomes in CU studies* – PFS or OS. *Study design* – cohort studies or randomized clinical trials. Studies were excluded if the study population excluded patients with any of the eight clinically relevant driver mutations; published in a non-English language; published before 2012 (one year before the first FDA approval of liquid biopsy test); and were review articles or conference proceedings.

Study identification

Relevant studies published between January 2012 and July 2023 were identified by searching MEDLINE and Embase databases with predefined search strategies (Supplementary File, [Table S2](#)) through the Embase platform. Furthermore, the official websites of three commonly used ctDNA tests (Guardant360[®] CDx, InvisionFirst[®]-Lung, and FoundationOne[®] Liquid CDx) as well as reference lists of included studies and previous systematic literature reviews¹⁸⁻²⁶ were searched for additional potentially eligible studies.

Study selection

Two reviewers (CC and MD) screened the identified abstracts using an open-source, active learning software/platform, ASReview, following recommended approaches for automated screening.³⁰⁻³⁴ Studies identified as eligible during abstract screening were subsequently screened at a full-text stage by the same two reviewers according to the eligibility criteria to determine the final set of included studies. Following reconciliation between the two investigators, a third reviewer (JJ) was included to reach a consensus for

any remaining discrepancies. The process of study identification and selection was summarized with a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

Data extraction

Two reviewers (CC and MD) extracted data on study characteristics, test and intervention characteristics, patient characteristics, and outcomes for the final list of included studies. Data was stored and managed in a Microsoft Excel workbook. *Trial characteristics* - author name, publication year, country, sample size, oncogenic driver mutations of interest, study duration, and patient in/exclusion criteria; *patient characteristics* - disease stage, smoking status, race/ethnicity, gender, and age; *test and intervention characteristics* - ctDNA and NGS technologies and in addition for CU studies, targeted treatment regimen; *Outcomes in CV studies* – overall and oncogenic driver mutation-specific sensitivity, specificity, TP, FP, FN, TN; *Outcomes in CU studies* - PFS and/or OS.

Quality assessment

Two authors (CC and MD) independently assessed the quality of included CV studies based on the revised Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) criteria.³⁵ Discrepancies were resolved by consensus.

Analysis

Study-specific, sample (mutation)-level TP, FP, FN, and TN frequency data for oncogenic driver mutations were used to estimate the overall and mutation-specific sensitivity and specificity for each study as well as across studies by meta-analyses. The overall sensitivity and specificity were estimated based on studies that simultaneously tested at least four driver mutations. Study-specific sensitivity and specificity estimates along with 95% confidence intervals (95% CIs) were obtained according to the Wilson method using continuity corrected cell counts.^{36,37} The bivariate random-effects model proposed by Reitsma et al. (2005) utilizes the standard frequentist approach and was used in our study to obtain pooled estimates for sensitivity and specificity.³⁸⁻⁴⁰ Meta-analysis results were also presented with

summary receiver operating characteristics (SROC) curves and 95% confidence regions for the pooled sensitivity and specificity estimates. In addition, as a sensitivity analysis, a Bayesian bivariate random-effects meta-analysis was performed to avoid normal approximations of the likelihood and to obtain predictive distributions of the sensitivity and specificity to predict results in a new study. Both the frequentist and Bayesian approaches used in our study modeled the sum and differences of true positive and false positive rates as random effects. The Bayesian receiver operating characteristics (BSROC) curves and Bayesian area under the BSROC curve (BAUC) were reported with 95% CI for the ctDNA detection of any and each driver mutation. An AUC of 0.7 or higher is generally considered good accuracy and 0.6-0.7 is considered sufficient; an AUC below 0.5 indicates the test is not useful.⁴¹ There is no standard for good sensitivity/specificity of DNA testing. A ctDNA test might be considered acceptable for coverage if its performance is similar to the FDA-approved Guardant360[®].⁴² All analyses were conducted using RStudio, version 4.1.2 (©2009-2022 RStudio, PBC) using packages “mada” and “bamdit”.^{38,43} Progression-free/overall survival (PFS/OS) was summarized for CU studies.

Results

Study selection

Our initial search generated 1,749 potentially relevant publications. After screening titles and abstracts, 58 publications were selected for full-text review. A total of 20 publications corresponding to 20 studies were selected for inclusion ([Figure 1](#)). Sixteen studies⁴⁴⁻⁵⁹ reported CV only, one study⁶⁰ reported CU only, and one study⁶¹ provided information on both CV and CU.

Clinical validity of ctDNA testing to identify oncogenic driver mutations

Characteristics of included studies

The majority of the included CV studies^{45–47,49–51,53–59,61,62} used a prospective cohort design (83.3%, 15/18), and three studies^{44,48,52} used retrospective medical record data. All studies evaluated ctDNA-based NGS technologies, with eleven studies^{44–46,48–52,58,61,62} evaluating branded ctDNA tests and seven studies^{47,53–57,59} describing the ctDNA technologies without a specific brand name (Table 1).

Guardant360[®] was the most used ctDNA technology (33.3%, 6/18). TB was the reference standard in all included studies.

Seven studies^{45,46,50,51,55,56,61} included only untreated patients, five studies^{44,47,48,53,58} included 24–92% of untreated patients, and in six studies^{49,52,54,57,59,62} the proportion of untreated patients was unclear. Among studies reporting patient racial/ethnic information^{46,51,52,61}, most patients were Caucasian (Table S3). The type of driver mutations examined differed across studies. No study reported CV on all eight clinically relevant driver mutations. Most of the studies evaluated ctDNA detection of EGFR (77.8%, 14/18) and *KRAS* (61.1%, 11/18). Thirteen studies reported CV information on six or more mutations (Table S4).

Four studies were found to have no risk of bias or applicability concerns and nine studies had two or fewer items of concern according to the QUADAS-2 instrument (Table S16). Eight studies were found to have unclear-risk items, indicating potential issues with the reporting of CV evidence.

Overall sensitivity and specificity for detection of any guideline-recommended oncogenic driver mutation

Overall sensitivity for the detection of any driver mutation varied between 0.52 and 0.81 across 13 studies that simultaneously detected four or more driver mutations (Figure 2 and Supplementary File Table S4).

The overall specificity varied between 0.88 and 1, with 11 studies reported a specificity value of 0.90 or higher. There was a small degree of between-study heterogeneity regarding sensitivity and specificity ($I^2=20\%$).⁶³ The pooled sensitivity obtained with meta-analysis using the frequentist approach was 0.69

(95% CI, 0.63–0.74) and the pooled specificity was 0.99 (95% CI, 0.97–1.00). Bayesian meta-analysis generated similar pooled estimates: sensitivity=0.70 (95% CI, 0.60-0.79); specificity=0.99 (95% CI, 0.97-1.00). The SROC curves with both approaches and the 95% confidence regions for the pooled sensitivity and specificity estimates are in [Figure 3](#). The BAUC was 0.71 (95% CI, 0.68-0.73). In general, overall sensitivity was higher in studies that used branded ctDNA tests compared to studies that did not describe tests with a specific brand name, except for the ResBio ctDx_Lung used in the study by Sabari et al. Overall specificity was high regardless of the ctDNA test used. We did not observe a positive relationship between sensitivity/specificity and study quality.

Sensitivity and specificity by oncogenic driver mutation

An overview of meta-analysis results of sensitivity and specificity by oncogenic driver mutation is presented in [Table 2](#). Corresponding SROCs and study-specific estimates are provided in Supplementary File, [Table S6-S12](#) and [Figure S2-S22](#). Sensitivity and specificity for detecting *EGFR* with ctDNA were reported in 14 studies.^{46,47,49–58,61,62} Study-specific sensitivity estimates varied between 0.56 and 0.83, and specificity estimates varied between 0.68 and 1 ([Figure S2](#)). The pooled sensitivity and specificity as estimated with the meta-analysis were 0.68 and 0.98 respectively, with a BAUC of 0.71 (95% CI, 0.68-0.73). Pooled sensitivity for the other driver mutations was generally less than 0.70 (except 0.77 for *KRAS*), but specificity was close to 1. The pooled sensitivity for each driver mutation detection was generally higher for branded ctDNA tests compared to tests that were not described with a specific brand name; overall pooled specificity was still high regardless of ctDNA tests used. We also conducted an analysis of sensitivity and specificity of ctDNA testing for detecting different mutation classes, including SNVs, indels and fusions. We found that only SNVs are associated with an acceptable BAUC of 0.72 (95% CI, 0.70-0.74), while the other mutation classes demonstrated a BAUC of around 0.52 ([Tables S13-S15](#), [Figures S23-S31](#)).

Clinical utility of ctDNA testing to inform 1L treatment in patients with aNSCLC

Three studies evaluated the CU of ctDNA testing to inform 1L targeted therapy.^{60,61,64} Madison *et al.* compared PFS among patients on matched 1L therapies following comprehensive genomic profiling of ctDNA (FoundationOne[®]Liquid or FoundationACT[®]) and TB (FoundationOne[®]CDx or FoundationOne[®]) using data from the Flatiron Health-Foundation Medicine Clinico-Genomic Database.⁶⁰ Median PFS in the ctDNA group (n=33) was 13.8 (95% CI, 8.9-NA) months and 10.6 (95% CI, 8.7-13.6) months in the TB group (n=229). The hazard ratio of ctDNA vs. TB regarding PFS was 0.68 (95% CI, 0.36-1.26) – not statistically significant. OS was only reported according to matched or unmatched 1L treatment but not specifically for TB and ctDNA. Palmero *et al.* reported PFS for 41 patients treated with 1L targeted therapy informed by ctDNA or TB testing (Guardant360[®]) (median, 8.6; 95% CI, 7.6-11.6 months).⁶¹ According to the reported Kaplan-Meier (KM) curves, PFS was comparable between these two groups. Jee *et al.* reported OS for ctDNA-matched (median OS of 39 months) and tissue-matched treatment-naïve patients (29 months), however they did not provide a relative treatment effect estimate independent of whether driver mutations were detected with ctDNA to infer the clinical utility of ctDNA relative to TB testing.⁶⁴

Discussion

Our meta-analyses showed that ctDNA testing demonstrated acceptable sensitivity and high specificity for detecting any guideline-recommended oncogenic driver mutation. However, the sensitivity of ctDNA testing varied widely by driver mutation. Pooled sensitivity estimates indicated acceptable performance for KRAS, but estimates were less than 70% for the other driver mutations. Specificity of ctDNA testing was high for all mutations. Evidence regarding the CU of ctDNA testing relative to TB was limited. According to a single study, there was no difference in PFS between ctDNA and TB tests followed by 1L targeted therapy.

To our knowledge, this is the first systematic literature review and meta-analysis specifically focused on the diagnostic performance of ctDNA testing for multi-gene detection with NGS in patients with aNSCLC to inform 1L targeted therapy. In contrast to previous systematic literature reviews/meta-analyses on ctDNA detection of EGFR or KRAS^{18,22,23,25,26}, we assessed the diagnostic accuracy of ctDNA for detecting multiple driver mutations simultaneously rather than focusing on single mutation thereby providing more relevant information for routine clinical practice where ctDNA-based NGS is rapidly being adopted. Meta-analysis of diagnostic accuracy can be performed based on results reported at the patient level, or the sample (mutation)-level. Previous meta-analyses mostly reported ctDNA performance at the patient level^{18,22,23} or did not specify the data level.^{19,20} We used sample-level data to facilitate diagnostic performance by driver mutation as well as mutation class, which allows for a more detailed assessment of the CV of ctDNA testing and more mutation-specific information.

In our systematic review, we included any study evaluating ctDNA testing for which the study population included at least 80% aNSCLC patients and any proportion of treatment-naïve patients, which did not strictly align with our target patient population of interest – aNSCLC patients initiating 1L treatment. The reason to cast a wider net was to ensure we did not miss any study that reported relevant subgroup results. In the actual meta-analyses of CV, we only included data for aNSCLC patients. Some of these studies included both treatment-naïve and treatment-experienced patients, however, we do not have reason to believe that this has (externally) biased the estimates of diagnostic accuracy focused on the presence of driver mutation for the 1L target population. If we excluded studies comprising not only treatment-naïve patients, the evidence base would have been limited.

The available evidence base regarding the CV and CU of ctDNA was limited. The CV studies had small sample sizes which made it challenging to reliably estimate sensitivity for rare driver mutations. Regarding the risk of bias, for only four studies we did not identify any concerns. Although we did not observe a difference in overall diagnostic performance between higher and lower quality studies, the limited evidence base makes interpretation of this observation difficult. Similarly, the limited number of

studies did not make it feasible to reliably evaluate potential drivers of between-study differences in ctDNA testing performance either.

Despite the limitations of the available evidence base, the CV findings from our study can have important clinical implications for the role of ctDNA testing in NSCLC. Our results support current guidelines that recommend ctDNA testing as a complementary approach rather than a replacement for TB to inform 1L therapy, unless there is insufficient tissue for NGS or risks of biopsy are excessive.^{7,65} This is particularly relevant for specific driver mutations such as *EGFR* and *KRAS*, which had much higher detection rates than *ROS1* or *ALK*. While multiple factors have been implicated in the sensitivity of ctDNA including tumor burden and the presence of osseous metastases⁶⁸, molecular heterogeneity has been relatively unexplored as a predictor of sensitivity for ctDNA. In particular, molecular fusions such as *ALK* and *ROS1* may be less detectable via ctDNA given the number of potential fusion partners which makes development of a sensitive ctDNA assay challenging.^{66,67} For example, in a patient who has never smoked where risks of biopsy may be high or the initial biopsy may not have sufficient tissue, a negative ctDNA test may not be sufficient to rule out the presence of a *ROS1* fusion which has profound implications for clinical decision-making around repeat biopsies at time of diagnosis and in the future. In addition to informing 1L therapy, ctDNA findings before initiating treatment will also help with the interpretation of ctDNA to monitor disease progression and tumor burden. For example, the absence of a mutation at follow-up that was detected upfront is indicative of effective treatment.⁶⁹ Ongoing presence of a particular mutation after initiation of therapy (even in the absence of a targetable mutation) may reflect a higher underlying tumor burden and an independent risk factor for survival.⁷⁰

There are only two studies that have evaluated the impact of the CU of ctDNA in terms of PFS in 1L aNSCLC patients, and no studies evaluating OS. This is a key area of future study, as baseline ctDNA may rarely pick up a mutation that would not be detected with tissue biopsy and influence treatment decisions, ultimately impacting OS. Additionally, we only included PFS and OS in our assessment of CU of ctDNA, and future studies may also want to examine other CU measures such as turn-around time of

test results, time to initiation of therapy, number of patients matched to targeted therapy, or treatment decision impacts.

Conclusion

Based on the currently available evidence, ctDNA testing in patients with aNSCLC has an overall acceptable diagnostic accuracy for detecting any guideline-recommended oncogenic driver mutation when using TB as the reference standard. However, the sensitivity of ctDNA testing varies greatly depending on the specific oncogenic driver mutation. At the time of this review, there were limited studies on less common driver mutations, and this is an essential area of future research. Given the current detection rates, ctDNA cannot be recommended as a replacement for TB, considering the actionable driver mutations of interest, unless there is insufficient tissue for TB or risks associated with the procedure. As the technologies around ctDNA testing and NGS analysis continue to evolve, we anticipate new studies will become available at a rapid pace, necessitating timely updates to this systematic review and meta-analysis.

ACKNOWLEDGEMENT

Author Contributions:

Concept and design: Jansen, Chen.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Chen, Jansen, Ragavan

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Chen, Jansen

Administrative, technical, or material support: Douglas, Phillips

Supervision: Jansen, Phillips

Conflict of Interest Disclosures: Mr. Douglas has disclosed that he receives consulting fees from Illumina, Inc. Dr. Phillips has disclosed that she receives consulting fees from Illumina, Inc; serves on an evidence review panel (the California Technology Assessment Forum), which is an independent appraisal committee for the Institute for Clinical and Economic Review (ICER); served as an advisor on two panels for Roche Diagnostics on tumor sequencing; and served on the Association of Community Cancer Centers Policy Advisory Committee in 2021. Dr. Jansen is a part-time salaried employee of the Precision Medicine Group, where he provides methodological expertise for health economics and outcomes research studies unrelated to this paper. Dr. Ragavan serves as a consultant/advisor to Trial Library and received royalties from Pharmacy Times. CC has nothing to disclose.

REFERENCES

1. Non-Small Cell Lung Cancer Treatment (PDQ®)—Health Professional Version - NCI. Published September 30, 2022. Accessed October 4, 2022. <https://www.cancer.gov/types/lung/hp/non-small-cell-lung-treatment-pdq>
2. Lung Cancer - Non-Small Cell - Statistics. Cancer.Net. Published June 25, 2012. Accessed September 9, 2022. <https://www.cancer.net/cancer-types/lung-cancer-non-small-cell/statistics>
3. What Is Lung Cancer? | Types of Lung Cancer. Accessed September 9, 2022. <https://www.cancer.org/cancer/lung-cancer/about/what-is.html>
4. Majeed U, Manochakian R, Zhao Y, Lou Y. Targeted therapy in advanced non-small cell lung cancer: current advances and future trends. *Journal of Hematology & Oncology*. 2021;14(1):108. doi:10.1186/s13045-021-01121-2
5. Vestergaard HH, Christensen MR, Lassen UN. A systematic review of targeted agents for non-small cell lung cancer. *Acta Oncologica*. 2018;57(2):176-186. doi:10.1080/0284186X.2017.1404634
6. Yuan M, Huang LL, Chen JH, Wu J, Xu Q. The emerging treatment landscape of targeted therapy in non-small-cell lung cancer. *Sig Transduct Target Ther*. 2019;4(1):1-14. doi:10.1038/s41392-019-0099-9
7. Ettinger DS, Wood DE, Aisner DL, et al. Non–Small Cell Lung Cancer, Version 3.2022, NCCN Clinical Practice Guidelines in Oncology. *Journal of the National Comprehensive Cancer Network*. 2022;20(5):497-530. doi:10.6004/jnccn.2022.0025
8. De Maglio G, Pasello G, Dono M, et al. The storm of NGS in NSCLC diagnostic-therapeutic pathway: How to sun the real clinical practice. *Critical Reviews in Oncology/Hematology*. 2022;169:103561. doi:10.1016/j.critrevonc.2021.103561
9. Pennell NA, Mutebi A, Zhou ZY, et al. Economic Impact of Next-Generation Sequencing Versus Single-Gene Testing to Detect Genomic Alterations in Metastatic Non–Small-Cell Lung Cancer Using a Decision Analytic Model. *JCO Precision Oncology*. 2019;(3):1-9. doi:10.1200/PO.18.00356
10. Aggarwal C, Rolfo CD, Oxnard GR, Gray JE, Sholl LM, Gandara DR. Strategies for the successful implementation of plasma-based NSCLC genotyping in clinical practice. *Nat Rev Clin Oncol*. 2021;18(1):56-62. doi:10.1038/s41571-020-0423-x
11. Rolfo C, Mack PC, Scagliotti GV, et al. Liquid Biopsy for Advanced Non-Small Cell Lung Cancer (NSCLC): A Statement Paper from the IASLC. *Journal of Thoracic Oncology*. 2018;13(9):1248-1268. doi:10.1016/j.jtho.2018.05.030
12. Commissioner O of the. FDA Approves First Liquid Biopsy Next-Generation Sequencing Companion Diagnostic Test. FDA. Published August 11, 2020. Accessed October 4, 2022. <https://www.fda.gov/news-events/press-announcements/fda-approves-first-liquid-biopsy-next-generation-sequencing-companion-diagnostic-test>
13. U.S. Food and Drug Administration (FDA) Approves FoundationOne®LiquidCDx as a Companion Diagnostic for Pfizer’s BRAFTOVI® (encorafenib) in Combination With Cetuximab to Identify Patients With BRAF V600E Alterations in Metastatic Colorectal Cancer | Foundation Medicine.

Accessed October 16, 2023. <https://www.foundationmedicine.com/press-releases/f9b285eb-db6d-4f61-856c-3f1edb803937>

14. Chae YK, Davis AA, Carneiro BA, et al. Concordance between genomic alterations assessed by next-generation sequencing in tumor tissue or circulating cell-free DNA. *Oncotarget*. 2016;7(40):65364-65373. doi:10.18632/oncotarget.11692
15. Weber S, Spiegl B, Perakis SO, et al. Technical Evaluation of Commercial Mutation Analysis Platforms and Reference Materials for Liquid Biopsy Profiling. *Cancers (Basel)*. 2020;12(6):1588. doi:10.3390/cancers12061588
16. Burke W. Genetic Tests: Clinical Validity and Clinical Utility. *Curr Protoc Hum Genet*. 2014;81:9.15.1-9.15.8. doi:10.1002/0471142905.hg0915s81
17. Promoting Safe and Effective Genetic Testing in the United States. Genome.gov. Accessed November 4, 2022. <https://www.genome.gov/10001733/genetic-testing-report>
18. Passiglia F, Rizzo S, Di Maio M, et al. The diagnostic accuracy of circulating tumor DNA for the detection of EGFR-T790M mutation in NSCLC: a systematic review and meta-analysis. *Sci Rep*. 2018;8:13379. doi:10.1038/s41598-018-30780-4
19. van Delft F, Koffijberg H, Retèl V, van den Heuvel M, IJzerman M. The Validity and Predictive Value of Blood-Based Biomarkers in Prediction of Response in the Treatment of Metastatic Non-Small Cell Lung Cancer: A Systematic Review. *Cancers (Basel)*. 2020;12(5):1120. doi:10.3390/cancers12051120
20. Sebastião MM, Ho RS, de Carvalho JPV, Nussbaum M. Diagnostic Accuracy of Next Generation Sequencing Panel using Circulating Tumor DNA in Patients with Advanced Non-Small Cell Lung Cancer: A Systematic Review and Meta-Analysis. *J Health Econ Outcomes Res*. 2020;7(2):158-163. doi:10.36469/jheor.2020.17088
21. Esagian SM, Grigoriadou GI, Nikas IP, et al. Comparison of liquid-based to tissue-based biopsy analysis by targeted next generation sequencing in advanced non-small cell lung cancer: a comprehensive systematic review. *J Cancer Res Clin Oncol*. 2020;146(8):2051-2066. doi:10.1007/s00432-020-03267-x
22. Luo J, Shen L, Zheng D. Diagnostic value of circulating free DNA for the detection of EGFR mutation status in NSCLC: a systematic review and meta-analysis. *Sci Rep*. 2014;4:6269. doi:10.1038/srep06269
23. Qiu M, Wang J, Xu Y, et al. Circulating Tumor DNA Is Effective for the Detection of EGFR Mutation in Non-Small Cell Lung Cancer: A Meta-analysis. *Cancer Epidemiology, Biomarkers & Prevention*. 2015;24(1):206-212. doi:10.1158/1055-9965.EPI-14-0895
24. Saarenheimo J, Andersen H, Eigeliene N, Jekunen A. Gene-Guided Treatment Decision-Making in Non-Small Cell Lung Cancer – A Systematic Review. *Front Oncol*. 2021;11:754427. doi:10.3389/fonc.2021.754427
25. Chen Z, Miao H, Zeng Q, Xu S, Chen Z, Liu K. Circulating cell-free DNA as a diagnostic and prognostic biomarker for non-small-cell lung cancer: a systematic review and meta-analysis. *Biomarkers in Medicine*. 2020;14(7):587-597. doi:10.2217/bmm-2018-0093

26. Wang N, Zhang X, Wang F, et al. The Diagnostic Accuracy of Liquid Biopsy in EGFR-Mutated NSCLC: A Systematic Review and Meta-Analysis of 40 Studies. *SLAS Technology*. 2021;26(1):42-54. doi:10.1177/2472630320939565
27. Dang DK, Park BH. Circulating tumor DNA: current challenges for clinical utility. *J Clin Invest*. 132(12):e154941. doi:10.1172/JCI154941
28. Stadler JC, Belloum Y, Deitert B, et al. Current and Future Clinical Applications of ctDNA in Immuno-Oncology. *Cancer Res*. 2022;82(3):349-358. doi:10.1158/0008-5472.CAN-21-1718
29. Walcott SE, Miller FA, Dunsmore K, Lazor T, Feldman BM, Hayeems RZ. Measuring clinical utility in the context of genetic testing: a scoping review. *Eur J Hum Genet*. 2021;29(3):378-386. doi:10.1038/s41431-020-00744-2
30. ASReview LAB: Active learning for Systematic Reviews — ASReview LAB documentation. Accessed August 23, 2022. <https://asreview.readthedocs.io/en/latest/index.html>
31. van de Schoot R, de Bruin J, Schram R, et al. An open source machine learning framework for efficient and transparent systematic reviews. *Nat Mach Intell*. 2021;3(2):125-133. doi:10.1038/s42256-020-00287-7
32. Ferdinands G, Schram R, de Bruin J, et al. *Active Learning for Screening Prioritization in Systematic Reviews - A Simulation Study*. Open Science Framework; 2020. doi:10.31219/osf.io/w6qbg
33. ASReview. Active Learning Explained. ASReview. Published April 12, 2022. Accessed August 23, 2022. <https://asreview.nl/blog/active-learning-explained/>
34. ASReview – Screen papers for literature review using machine learning. Singapore Management University (SMU). Accessed August 23, 2022. <https://library.smu.edu.sg/topics-insights/asreview-screen-papers-literature-review-using-machine-learning>
35. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med*. 2011;155(8):529-536. doi:10.7326/0003-4819-155-8-201110180-00009
36. Wilson EB. Probable Inference, the Law of Succession, and Statistical Inference. *Journal of the American Statistical Association*. 1927;22(158):209-212. doi:10.1080/01621459.1927.10502953
37. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat. Med*. Published online 1998.
38. Sousa-Pinto PD with contributions from B. mada: Meta-Analysis of Diagnostic Accuracy. Published online July 15, 2022. Accessed September 13, 2022. <https://CRAN.R-project.org/package=mada>
39. Holling H, Böhning W, Masoudi E, Böhning D, Sangnawakij P. Evaluation of a new version of I2 with emphasis on diagnostic problems. *Communications in Statistics - Simulation and Computation*. 2020;49(4):942-972. doi:10.1080/03610918.2018.1489553
40. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology*. 2005;58(10):982-990. doi:10.1016/j.jclinepi.2005.02.022

41. Šimundić AM. Measures of Diagnostic Accuracy: Basic Definitions. *EJIFCC*. 2009;19(4):203-211.
42. LCD - MolDX: Plasma-Based Genomic Profiling in Solid Tumors (L38065). Accessed December 9, 2022. <https://www.cms.gov/medicare-coverage-database/view/lcd.aspx?lcdid=38065&ver=9&keyword=guardant360&keywordType=starts&areaId=all&docType=NCA,CAL,NCD,MEDCAC,TA,MCD,6,3,5,1,F,P&contractOption=all&sortBy=relevance&bc=1>
43. Verde PE. bamdit: Bayesian Meta-Analysis of Diagnostic Test Data. Published online April 4, 2022. Accessed September 13, 2022. <https://CRAN.R-project.org/package=bamdit>
44. Bustamante Alvarez JG, Janse S, Owen DH, et al. Treatment of Non–Small-Cell Lung Cancer Based on Circulating Cell-Free DNA and Impact of Variation Allele Frequency. *Clinical Lung Cancer*. 2021;22(4):e519-e527. doi:10.1016/j.clcc.2020.11.007
45. Fernandes MGO, Cruz-Martins N, Souto Moura C, et al. Clinical Application of Next-Generation Sequencing of Plasma Cell-Free DNA for Genotyping Untreated Advanced Non-Small Cell Lung Cancer. *Cancers (Basel)*. 2021;13(11):2707. doi:10.3390/cancers13112707
46. Leighl NB, Page RD, Raymond VM, et al. Clinical Utility of Comprehensive Cell-free DNA Analysis to Identify Genomic Biomarkers in Patients with Newly Diagnosed Metastatic Non–small Cell Lung Cancer. *Clinical Cancer Research*. 2019;25(15):4691-4700. doi:10.1158/1078-0432.CCR-19-0624
47. Li BT, Janku F, Jung B, et al. Ultra-deep next-generation sequencing of plasma cell-free DNA in patients with advanced lung cancers: results from the Actionable Genome Consortium. *Ann Oncol*. 2019;30(4):597-603. doi:10.1093/annonc/mdz046
48. Lin LH, Allison DHR, Feng Y, et al. Comparison of solid tissue sequencing and liquid biopsy accuracy in identification of clinically relevant gene mutations and rearrangements in lung adenocarcinomas. *Mod Pathol*. 2021;34(12):2168-2174. doi:10.1038/s41379-021-00880-0
49. Müller JN, Falk M, Talwar J, et al. Concordance between Comprehensive Cancer Genome Profiling in Plasma and Tumor Specimens. *Journal of Thoracic Oncology*. 2017;12(10):1503-1511. doi:10.1016/j.jtho.2017.07.014
50. Park S, Olsen S, Ku BM, et al. High concordance of actionable genomic alterations identified between circulating tumor DNA–based and tissue-based next-generation sequencing testing in advanced non–small cell lung cancer: The Korean Lung Liquid Versus Invasive Biopsy Program. *Cancer*. 2021;127(16):3019-3028. doi:10.1002/cncr.33571
51. Pritchett MA, Camidge DR, Patel M, et al. Prospective Clinical Validation of the InVisionFirst-Lung Circulating Tumor DNA Assay for Molecular Profiling of Patients With Advanced Nonsquamous Non–Small-Cell Lung Cancer. *JCO Precision Oncology*. 2019;(3):1-15. doi:10.1200/PO.18.00299
52. Schwaederlé MC, Patel SP, Husain H, et al. Utility of genomic assessment of blood-derived circulating tumor DNA (ctDNA) in patients with advanced lung adenocarcinoma. *Clin Cancer Res*. 2017;23(17):5101-5111. doi:10.1158/1078-0432.CCR-16-2497
53. Yao Y, Liu J, Li L, et al. Detection of circulating tumor DNA in patients with advanced non-small cell lung cancer. *Oncotarget*. 2016;8(2):2130-2140. doi:10.18632/oncotarget.12883

54. Xu S, Lou F, Wu Y, et al. Circulating tumor DNA identified by targeted sequencing in advanced-stage non-small cell lung cancer patients. *Cancer Lett.* 2016;370(2):324-331. doi:10.1016/j.canlet.2015.11.005
55. Liu L, Liu H, Shao D, et al. Development and clinical validation of a circulating tumor DNA test for the identification of clinically actionable mutations in nonsmall cell lung cancer. *Genes, Chromosomes and Cancer.* 2018;57(4):211-220. doi:10.1002/gcc.22522
56. Couraud S, Vaca-Paniagua F, Villar S, et al. Noninvasive Diagnosis of Actionable Mutations by Deep Sequencing of Circulating Free DNA in Lung Cancer from Never-Smokers: A Proof-of-Concept Study from BioCAST/IFCT-1002. *Clinical Cancer Research.* 2014;20(17):4613-4624. doi:10.1158/1078-0432.CCR-13-3063
57. Chen Y, Han T, Zhou Y, Mao B, Zhuang W. Comparing the efficacy of targeted next-generation sequencing in the identification of somatic mutations in circulating tumor DNA from different stages of lung cancer. *Neoplasma.* 2019;66(4):652-660. doi:10.4149/neo_2018_181130N910
58. Remon J, Lacroix L, Jovelet C, et al. Real-World Utility of an Amplicon-Based Next-Generation Sequencing Liquid Biopsy for Broad Molecular Profiling in Patients With Advanced Non-Small-Cell Lung Cancer. *JCO Precis Oncol.* 2019;3:PO.18.00211. doi:10.1200/PO.18.00211
59. Guo N, Lou F, Ma Y, et al. Circulating tumor DNA detection in lung cancer patients before and after surgery. *Sci Rep.* 2016;6(1):33519. doi:10.1038/srep33519
60. Madison R, Schrock AB, Castellanos E, et al. Retrospective analysis of real-world data to determine clinical outcomes of patients with advanced non-small cell lung cancer following cell-free circulating tumor DNA genomic profiling. *Lung Cancer.* 2020;148:69-78. doi:10.1016/j.lungcan.2020.07.033
61. Palmero R, Taus A, Viteri S, et al. Biomarker Discovery and Outcomes for Comprehensive Cell-Free Circulating Tumor DNA Versus Standard-of-Care Tissue Testing in Advanced Non-Small-Cell Lung Cancer. *JCO Precision Oncology.* 2021;(5):93-102. doi:10.1200/PO.20.00241
62. Sabari JK, Offin M, Stephens D, et al. A Prospective Study of Circulating Tumor DNA to Guide Matched Targeted Therapy in Lung Cancers. *J Natl Cancer Inst.* 2018;111(6):575-583. doi:10.1093/jnci/djy156
63. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence— inconsistency. *Journal of Clinical Epidemiology.* 2011;64(12):1294-1302. doi:10.1016/j.jclinepi.2011.03.017
64. Jee J, Lebow ES, Yeh R, et al. Overall survival with circulating tumor DNA-guided therapy in advanced non-small-cell lung cancer. *Nat Med.* 2022;28(11):2353-2363. doi:10.1038/s41591-022-02047-z
65. Lindeman NI, Cagle PT, Aisner DL, et al. Updated Molecular Testing Guideline for the Selection of Lung Cancer Patients for Treatment With Targeted Tyrosine Kinase Inhibitors: Guideline From the College of American Pathologists, the International Association for the Study of Lung Cancer, and the Association for Molecular Pathology. *Journal of Thoracic Oncology.* 2018;13(3):323-358. doi:10.1016/j.jtho.2017.12.001

66. Pascual J, Attard G, Bidard FC, et al. ESMO recommendations on the use of circulating tumour DNA assays for patients with cancer: a report from the ESMO Precision Medicine Working Group. *Annals of Oncology*. 2022;33(8):750-768. doi:10.1016/j.annonc.2022.05.520
67. Zografos E, Dimitrakopoulos FI, Koutras A. Prognostic Value of Circulating Tumor DNA (ctDNA) in Oncogene-Driven NSCLC: Current Knowledge and Future Perspectives. *Cancers (Basel)*. 2022;14(19):4954. doi:10.3390/cancers14194954
68. Lee Y, Park S, Kim WS, et al. Correlation between progression-free survival, tumor burden, and circulating tumor DNA in the initial diagnosis of advanced-stage EGFR-mutated non-small cell lung cancer. *Thorac Cancer*. 2018;9(9):1104-1110. doi:10.1111/1759-7714.12793
69. Singh AP, Li S, Cheng H. Circulating DNA in EGFR-mutated lung cancer. *Ann Transl Med*. 2017;5(18):379. doi:10.21037/atm.2017.07.10
70. Zou W, Yaung SJ, Fuhlbrück F, et al. ctDNA Predicts Overall Survival in Patients With NSCLC Treated With PD-L1 Blockade or With Chemotherapy. *JCO Precis Oncol*. 2021;5:827-838. doi:10.1200/PO.21.00057

Figure Legends

Figure 1. The PRISMA flowchart of included studies

Figure 2. Forest plot of sensitivity and specificity of ctDNA testing for multi-gene detection from bivariate random-effects meta-analyses (n=13)

Figure 3. Summary receiver operating characteristics (SROC) plots based on frequentist and Bayesian bivariate random-effects meta-analyses (n=13): (a) Each triangle identifies the true positive rate vs. the false positive rate (1- specificity) of each study (observed data). The black circle represents the summary estimate, and the solid contour shows the 95% confidence region around the summary estimate. The dotted contour indicates the 95% prediction region (the region within which a new study will lie); (b) The panel shows the Bayesian summary receiver operating characteristics (SROC) curve. Each blue circle indicates the true positive rate vs. the false positive rate (1- specificity) of each study, and different sizes are used for different sample sizes. The central line corresponds to the posterior median and the upper and lower curves correspond to the quantiles of the 2.5 % and 97.5 % respectively.

Table 1. Characteristics of included CV studies (n=18)

Author_Year	Country	No. Pts*	LB Technology	TB Technology	Time between testing (days)	Test genes/driver mutations (included in analysis)
Bustamante Alvarez_2021	US	143	Guardant360®	Fluorescence in situ hybridization (FISH)	≤84	<i>EGFR</i> , <i>ALK</i> fusions, <i>BRAF</i> , <i>MET</i> exon 14 skipping mutation, <i>ROS1</i> mutations and <i>RET</i> fusions; <i>KRAS</i>
Fernandes_2021	Portugal	115	Oncomine™ Lung cfDNA	Unspecified	--	Overall Oncomine panel (<i>EGFR</i> , <i>KRAS</i> , <i>ALK</i> , <i>BRAF</i> , <i>MET</i> , <i>ROS1</i> , <i>ERBB2</i> , <i>MAP2K1</i> , <i>NRAS</i> , <i>PIK3CA</i> , and <i>TP53</i>)
Leighl_2019	US	282	Guardant360®	Unspecified	--	<i>EGFR</i> , <i>ALK</i> fusion, <i>BRAF</i> V600E, <i>MET</i> amplification, <i>MET</i> exon 14 skipping, <i>ROS1</i> fusion, <i>RET</i> fusion, <i>ERBB2</i> mutation
Li_2019	US	127	Unspecified	MSK-IMPACT	207 (median)	<i>EGFR</i> , <i>KRAS</i> , <i>ALK</i> , <i>BRAF</i> , <i>MET</i> , <i>ROS1</i> , <i>RET</i> , <i>ERBB2</i>
Lin_2021	US	100	Guardant360®	NYU Oncomine™ Focus	12 (median)	Overall Guardant 360 panel (patient level)
Müller_2017	Germany	98	NEOliquid®	NEOPlus, PCR and FISH	--	<i>EGFR</i> , <i>KRAS</i> , <i>ALK</i> , <i>BRAF</i> , <i>ROS1</i> , <i>TP53</i>
Palmero_2021	Spain	186	Guardant360®	Unspecified	--	<i>EGFR</i> , <i>ALK</i>
Park_2021	Korea	421	Guardant360®	Oncomine™ Focus	≤1	<i>EGFR</i> Del 9, <i>EGFR</i> L858R, <i>KRAS</i> , <i>ALK</i> , <i>BRAF</i> V600E, <i>MET</i> exon 14, <i>MET</i> amplification, <i>ROS1</i> , <i>RET</i> , <i>ERBB2</i> exon 20
Pritchett_2019	US	264	InVisionFirst®	Custom-designed SureSelect XT assay and FusionPlex Tumor Panel	≤84	<i>EGFR</i> (exons 18-21), <i>KRAS</i> , <i>ALK/ROS1</i> , <i>BRAF</i> V600E, <i>MET</i> exon 14, <i>ERBB2</i> exon 20, <i>STK11</i>
Schwaederlé_2017	US	88	Guardant360®	FoundationOne	--	Overall Guardant 360 panel (patient level)
Yao_2017	China	39	Unspecified	Unspecified	≤7	<i>EGFR</i> , <i>KRAS</i> , <i>ALK</i> , <i>RET</i> , <i>PIK3CA</i>
Xu_2016	China	42	Unspecified	Unspecified	--	50 gene panel (<i>EGFR</i> , <i>KRAS</i> , <i>BRAF</i> , <i>ERBB2</i> , <i>PIK3CA</i> , <i>TP53</i> identified)
Liu_2018	China	72	Unspecified	Unspecified	≤7	<i>EGFR</i> L858R, <i>EGFR</i> exon 19 del, <i>KRAS</i> G12X, <i>ALK</i> rearrangement

Couraud_2014	France	106	Unspecified	Unspecified	--	<i>EGFR</i> (exons 18-21), <i>KRAS</i> (exons 2 & 3), <i>BRAF</i> (exons 11 & 15), <i>ERBB2</i> (exons 19 & 20), <i>PI3KCA</i> (exons 9 & 20)
Chen_2019	China	50	Unspecified	Unspecified	--	>200 gene variants; specific CV on <i>EGFR</i>
Remon_2019	France	194	InVisionFirst®	Unspecified	34 (median)	<i>EGFR</i> (exons 18-21), <i>KRAS</i> , <i>BRAF</i> V600E, <i>MET</i> exon 14, <i>ERBB2</i> exon 20, <i>STK11</i>
Guo_2016	China	41	SV-CA50-ctDNA panel	Ion Proton System, Ion PI sequencing	≤13	<i>EGFR</i> , <i>KRAS</i> , <i>BRAF</i> , <i>ERBB2</i> , <i>PIK3CA</i> , <i>TP53</i>
Sabari_2019	US and Australia	210	ResBio ctDx-Lung	MSK-IMPACT	21 (median)	<i>EGFR</i> , <i>KRAS</i> , <i>ALK</i> , <i>BRAF</i> , <i>MET</i> , <i>ROS1</i> , <i>RET</i>

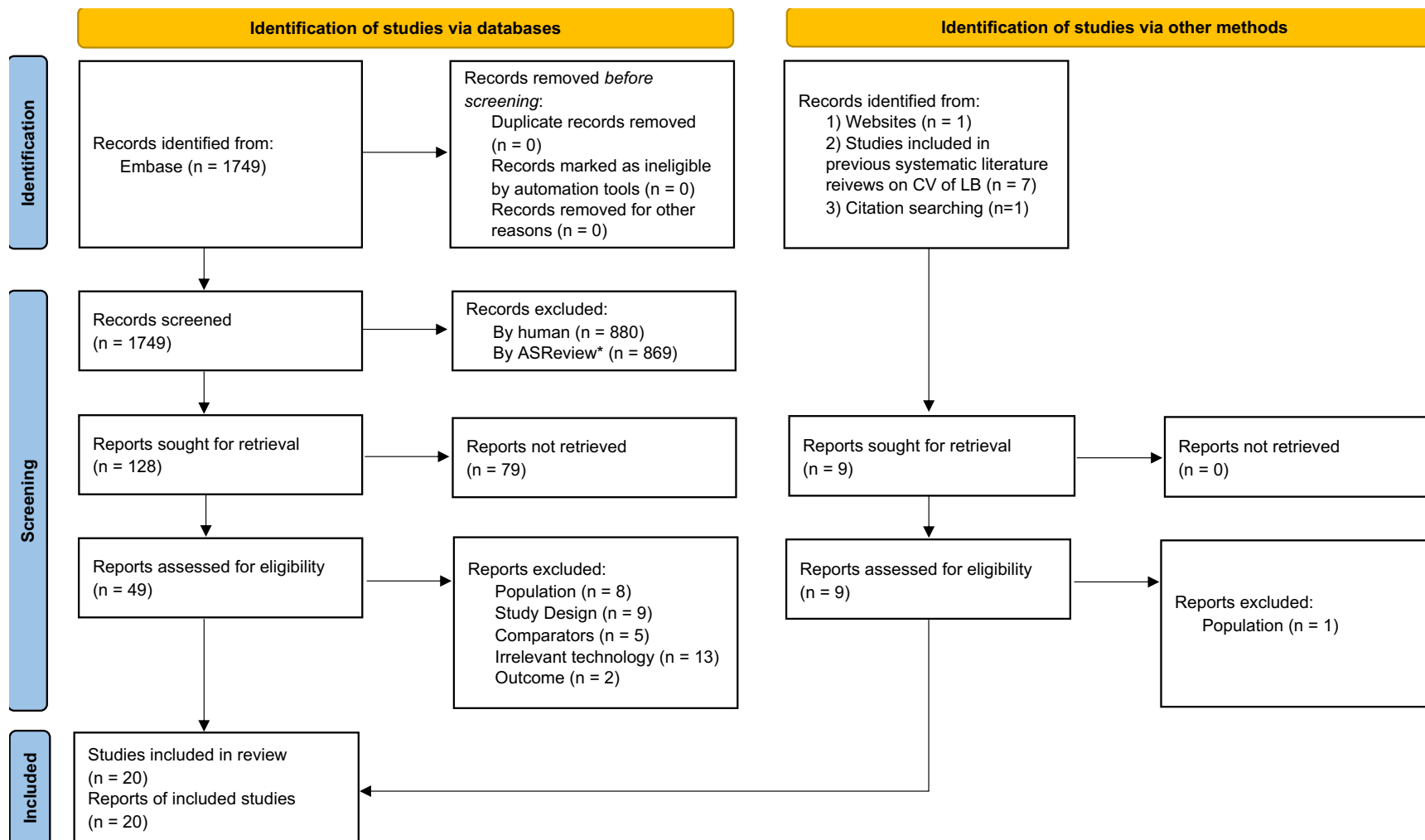
* Baseline overall patient sample size

Table 2. Results from frequentist bivariate random-effects meta-analyses of ctDNA testing for single-gene detection

	No. of studies	No. of Samples	Sensitivity (95% CI)	Specificity (95% CI)
<i>EGFR</i>	14	2047	0.68 (0.62, 0.74)	0.98 (0.95, 0.99)
<i>KRAS</i>	11	1067	0.77 (0.63, 0.86)	0.96 (0.90, 0.98)
<i>BRAF</i>	8	920	0.60 (0.39, 0.77)	0.99 (0.99, 1.00)
<i>ALK</i>	6	889	0.59 (0.45, 0.73)	0.99 (0.98, 1.00)
<i>ROS1</i>	4	631	0.29 (0.13, 0.53)	0.99 (0.97, 1.00)
<i>MET</i>	5	927	0.47 (0.33, 0.62)	0.98 (0.97, 0.99)
<i>RET</i>	3	478	0.38 (0.21, 0.58)	0.99 (0.97, 1.00)

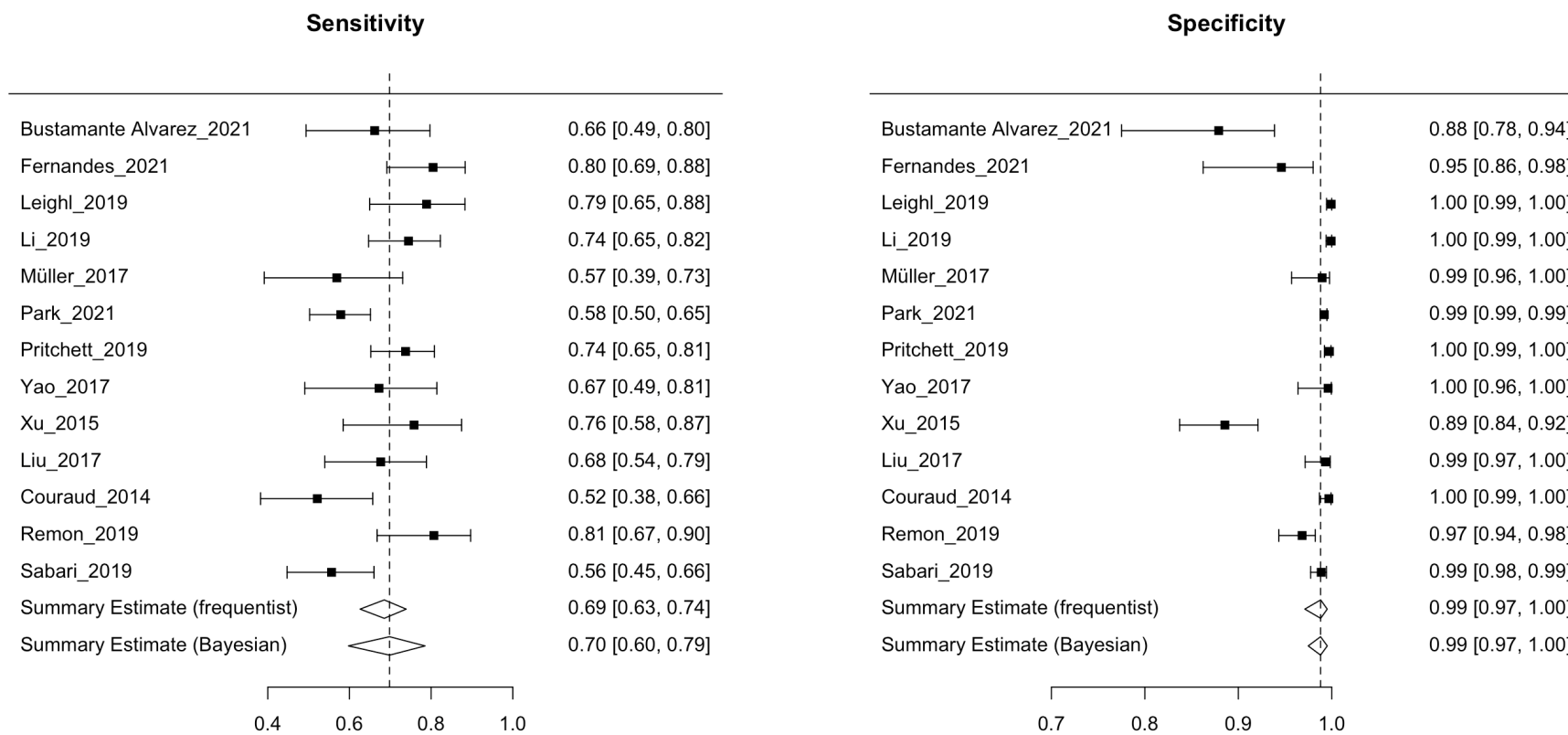
Note: Studies with 0 positive cases (based on the reference standard – tissue biopsy) were excluded from analyses.

Figure 1. The PRISMA flowchart of included studies



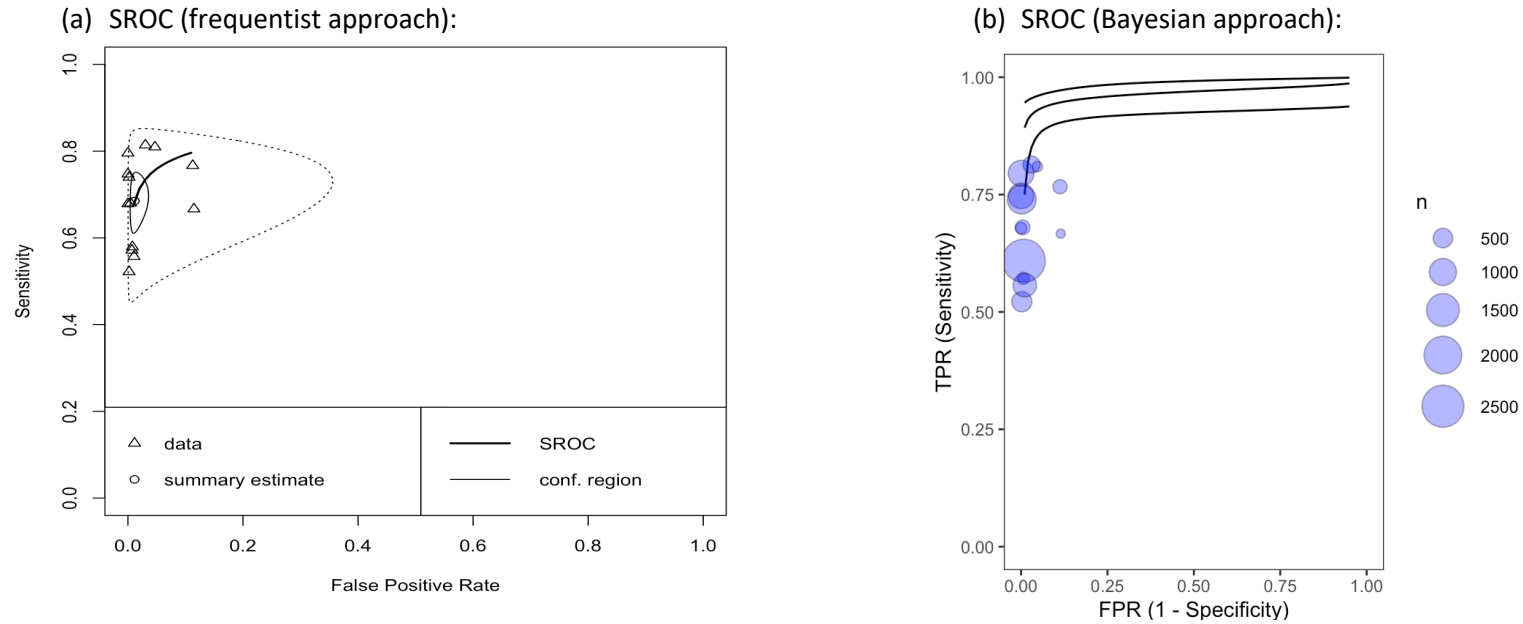
* ASReview is an open-source, active-learning software for screening titles and abstracts for systematic reviews.

Figure 2. Forest plot of sensitivity and specificity of ctDNA testing for multi-gene detection from bivariate random effects meta-analyses (n=13)



Note: Sensitivity and specificity of single studies were based on frequentist estimations. $I^2 = 20\%$ based on the frequentist approach

Figure 3. Summary receiver operating characteristics (SROC) plots based on frequentist and Bayesian bivariate random-effects meta-analyses (n=13)



(a) Each triangle identifies the true positive rate vs. the false positive rate (1- specificity) of each study (observed data). The black circle represents the summary estimate, and the solid contour shows the 95% confidence region around the summary estimate. The dotted contour indicates the 95% prediction region (the region within which a new study will lie). (b) The panel shows the Bayesian summary receiver operating characteristics (SROC) curve. Each blue circle indicates the true positive rate vs. the false positive rate (1- specificity) of each study, and different sizes are used for different sample sizes. The central line corresponds to the posterior median and the upper and lower curves correspond to the quantiles of the 2.5 % and 97.5 % respectively.