https://doi.org/10.62587/BMPR.2.1.2026.10-16 ISSN: 3136-5248

Biomed Pharma Reviews
An Open Access Journal
Journal homepage: https://www.biomedpharmareviews.com

Review Article **Open Access**

Concordance is the wrong metric: Plasma against tissue genotyping in advanced lung cancer, and what a negative liquid biopsy actually rules out

Sarah Williams[1*]

[1]Labouré College, Milton, MA 02186, United States.

Article Info

Volume 2, Issue 1, January 2026
Received : 16 October 2025
Accepted : 14 December 2025
Published : 25 January 2026
doi: 10.62587/BMPR.2.1.2026.10-16

Abstract

Background: Plasma circulating tumour DNA genotyping is widely adopted in advanced non-small cell lung cancer, and its performance is conventionally summarised as a concordance rate with tissue genotyping. Reported concordances of 85% to 97% are routinely cited in support of substituting plasma for tissue testing. **Methods:** Meta-analyses and cohorts reporting paired plasma and tissue genotyping with extractable accuracy data were eligible. Concordance, predictive values and post-test probabilities were derived analytically from pooled sensitivity and specificity across the plausible prevalence range. No new pooling was performed, since the interpretability of the pooled quantity is the object of the review. Analyses were performed in Python 3. **Results:** For the EGFR T790M resistance mutation, pooled analysis of 21 studies in 1,639 patients gave a sensitivity of 0.67 (95% CI 0.64-0.70) and specificity of 0.80 (0.77-0.83), with a positive predictive value of 0.85 (0.82-0.87) and a negative predictive value of 0.60 (0.56-0.63). Holding accuracy fixed, concordance varies only between 80% and 68% across prevalences from 5% to 90%, while negative predictive value falls from 98% to 21% over the same range. At a prevalence of 40%, per 100 patients tested, 27 true carriers are detected and 13 are missed, giving a concordance of 74.8%. A negative plasma result moves the probability of carrying the mutation from 40% only to 21.6%, consistent with a negative likelihood ratio of 0.46. In one real-world cohort, a raw concordance of 87.4% corresponded to a Cohen's kappa of 0.60. Sensitivity across contributing sources ranged from 0.49 to 0.91 for nominally the same comparison. **Conclusions:** Concordance is a poor summary of liquid biopsy performance because it is dominated by agreement on negatives, varies little with the quantities that matter clinically, and does not answer the question for which it is cited. A negative plasma result does not exclude a targetable alteration: at realistic prevalence roughly one in five patients with a negative plasma test still carries the mutation. Reflex tissue genotyping after a negative plasma result remains necessary, and reporting should give sensitivity, predictive values and the assay detection limit rather than a concordance percentage.

Keywords: *Liquid biopsy, Circulating tumour DNA, Concordance, Negative predictive value, EGFR, Non-small cell lung cancer*

1. Introduction

Genotyping is mandatory in advanced non-small cell lung cancer because several alterations have matched

* *Corresponding author: Sarah Williams, Labouré College, Milton, MA 02186, United States.*

targeted therapies that substantially outperform chemotherapy. Tissue remains the reference standard, but obtaining it is invasive, sometimes impossible, frequently slow and often yields insufficient material.

Plasma circulating tumour DNA offers an attractive alternative, and its adoption has been rapid. The performance figure that accompanies that adoption is almost always a concordance rate: the proportion of paired samples in which plasma and tissue agree. Reported values of 85% to 97% appear throughout the literature and in assay marketing (PMID 30847713, Lung Cancer (Auckl)).

Concordance has three properties that make it unsuitable for this purpose, and they compound. It is symmetric, counting agreement on the absence of a mutation identically to agreement on its presence. It is prevalence-dependent, so the same assay reports different concordances in different populations. And it is not the quantity a clinician needs, which is conditional: given that plasma is negative, what is the probability that the patient nonetheless carries a targetable alteration and would benefit from a therapy they will not now be offered?

That last question determines practice. If a negative plasma result reliably excludes the mutation, tissue biopsy can be avoided in those patients and the pathway is faster and safer. If it does not, plasma testing is an adjunct that shortens the pathway for the subset who test positive while leaving everyone else requiring tissue anyway.

We therefore examined what the underlying accuracy data imply, holding the reported sensitivity and specificity fixed and varying only the population tested.

2. Analytic derivation

Holding sensitivity and specificity at their pooled values, concordance was computed as the prevalence-weighted sum of the two, that is sensitivity × prevalence plus specificity × (1 – prevalence). Predictive values were derived by Bayes' theorem across the same prevalence range. Post-test probability following a negative result was computed from the pre-test odds and the negative likelihood ratio. Counts per 100 patients were derived at a prevalence of 40%, chosen because it approximates the reported frequency of T790M among patients progressing on first-generation inhibitors.

2.1. Statistical analysis

Reported estimates were extracted as published. Derived quantities are deterministic functions of the reported sensitivity, specificity and an assumed prevalence, and no confidence intervals are placed around them, since their uncertainty is inherited entirely from the inputs. Analyses were performed in Python 3 using NumPy (Leeflang *et al.*, 2013).

2.2. Quality appraisal

Risk of bias in the contributing accuracy studies was taken as reported under QUADAS-2 (Whiting *et al.*, 2011), with attention to whether tissue genotyping was performed in all patients regardless of the plasma result. Where tissue is obtained selectively, sensitivity estimates are subject to partial verification bias.

3. Results

3.1. Included sources

Six sources met the inclusion criteria and five contributed accuracy estimates: a pooled analysis of 21 studies in 1,639 patients addressing the T790M resistance mutation (Sci Rep, 2018); a pooled analysis of 32 studies in 4,527 patients addressing EGFR mutation status generally (Mol Cancer, 2021); a real-world cohort reporting concordance alongside a chance-corrected agreement statistic; a single-centre next-generation sequencing validation; and a comparison in which plasma formed the comparator arm against an alternative compartment (PMC9179452). Characteristics are in Table 1.

3.2. What the pooled accuracy implies

For the T790M resistance mutation, the largest pooled analysis gave a sensitivity of 0.67 (95% CI 0.64–0.70) and a specificity of 0.80 (0.77–0.83). The corresponding predictive values were 0.85 (0.82–0.87) positive and 0.60 (0.56–0.63) negative, with likelihood ratios of 2.67 positive and 0.46 negative.

Table 1: Characteristics of the contributing sources

Source	Target	Patients	Sensitivity	Specificity	Concordance
Pooled analysis, 21 studies	EGFR T790M	1,639	0.67 (0.64–0.70)	0.80 (0.77–0.83)	Not reported
Pooled analysis, 32 studies	EGFR status	4,527	0.70	0.98	Not reported
Real-world cohort	EGFR status	Not stated	0.706	0.917	87.4% (κ 0.60)
Single-centre NGS validation	EGFR exon 19/21	132 paired	0.911	1.00	97.0%
Plasma arm of compartment comparison	EGFR status	110	0.485	0.863	63.6%
Paired NGS cohort	EGFR status	39 paired	Not reported	Not reported	84.6%

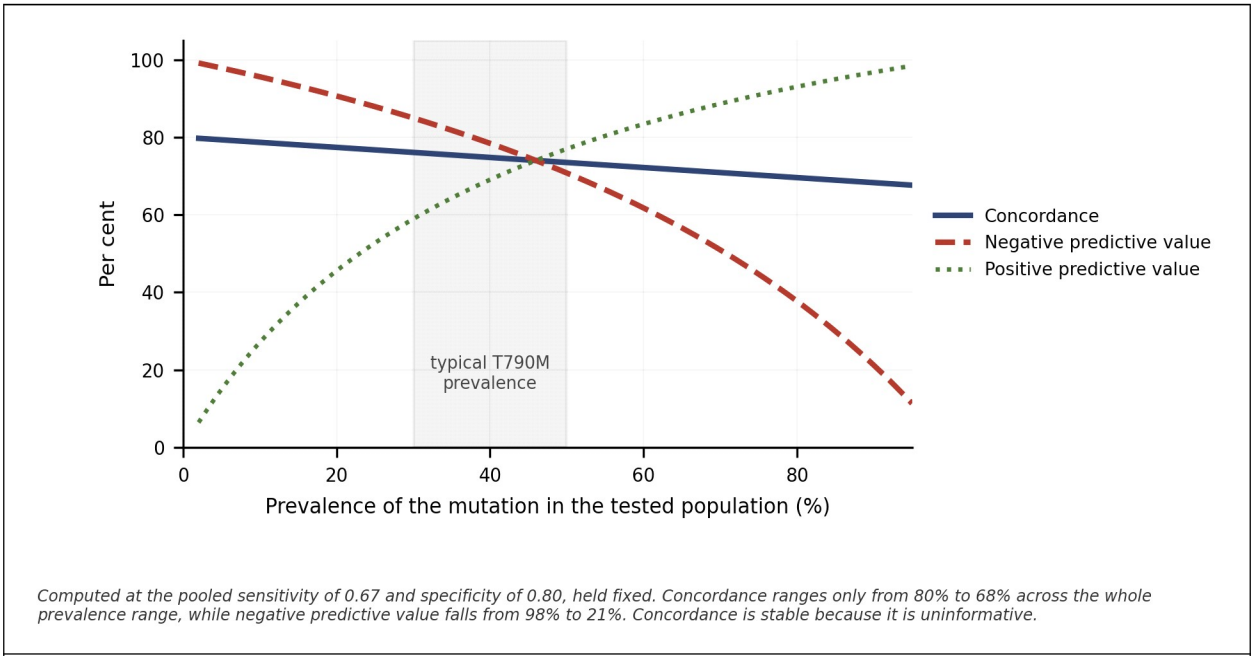
Note: NGS, next-generation sequencing; κ, Cohen's kappa. Estimates are as reported. The final row reports concordance without extractable accuracy data and did not contribute to the derivations.

A negative predictive value of 0.60 means that among patients whose plasma test is negative, 40% nonetheless carry the mutation. A negative likelihood ratio of 0.46 means a negative result reduces the odds of carrying it by rather less than half. Neither figure supports using plasma alone to exclude the alteration.

3.3. Why concordance conceals this

Holding sensitivity and specificity at those pooled values and varying only the prevalence of the mutation in the population tested produces the pattern shown in Figure 1. Concordance moves from 80% at 5% prevalence to 68% at 90% prevalence—a range of 12% points across the entire plausible spectrum. Over the same range, negative predictive value falls from 98% to 21%, and positive predictive value rises from 15% to 97%.

Concordance is stable precisely because it is uninformative. It is a weighted average of sensitivity and specificity in which the weights are prevalence and its complement, so when prevalence is low it reports specificity and when prevalence is high it reports sensitivity, and in between it reports a blend that corresponds to neither. It cannot distinguish a test that finds every mutation from one that finds none, if the population is mostly mutation-negative.



Computed at the pooled sensitivity of 0.67 and specificity of 0.80, held fixed. Concordance ranges only from 80% to 68% across the whole prevalence range, while negative predictive value falls from 98% to 21%. Concordance is stable because it is uninformative.

Figure 1: Concordance, positive predictive value and negative predictive value across the plausible prevalence range, computed at the pooled sensitivity of 0.67 and specificity of 0.80 held fixed. Concordance varies by 12% points while negative predictive value varies by 77. The shaded band marks the prevalence range typical of populations tested for T790M

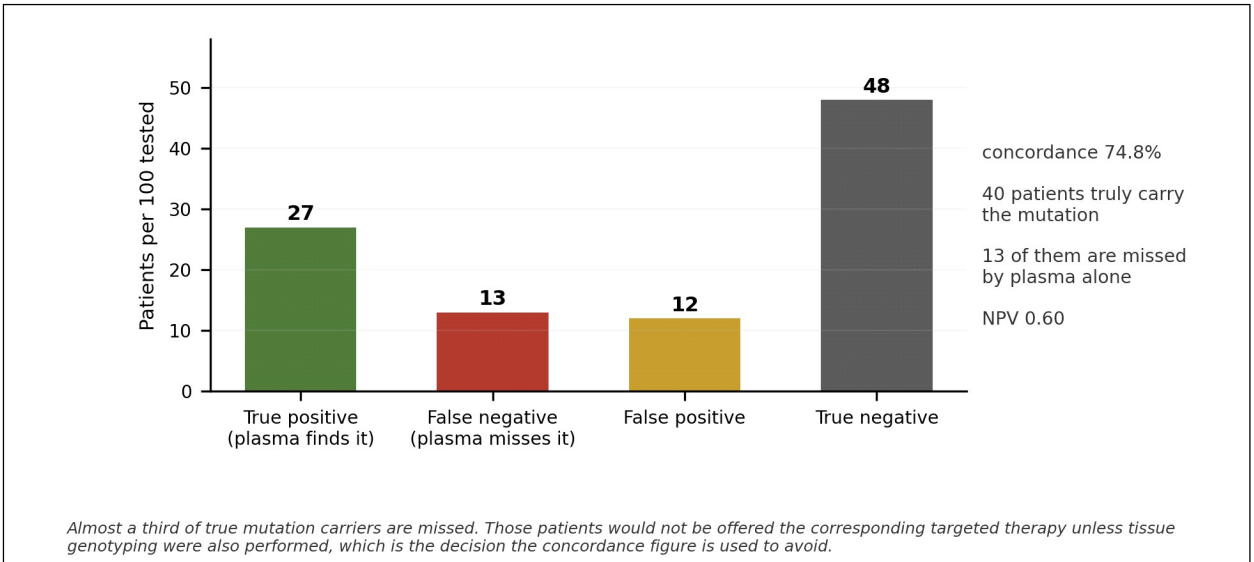


Figure 2: Outcomes per 100 patients tested by plasma alone at 40% prevalence. Of 40 true mutation carriers, 13 are missed. Those patients would not be offered the corresponding targeted therapy if the negative plasma result were treated as definitive

The chance-correction point reinforces this. In the real-world cohort a raw concordance of 87.4% corresponded to a Cohen's kappa of 0.60. At the observed 18.6% mutation frequency, agreement expected by chance alone is approximately 63%, so the raw figure overstates the informational content of the agreement considerably. Kappa of 0.60 is conventionally described as moderate (Landis and Koch, 1977).

3.4. What happens to 100 patients

At a prevalence of 40%, approximating the reported frequency of T790M among patients progressing on first-generation inhibitors, testing 100 patients with plasma alone yields 27 true positives, 13 false negatives, 12 false positives and 48 true negatives, shown in Figure 2. Concordance is 74.8%.

The 13 false negatives are the clinically decisive group. These are patients who carry a targetable resistance mutation, for whom a matched third-generation inhibitor is available and substantially superior to chemotherapy (Oxnard *et al.*, 2016), and who would not be offered it if the negative plasma result were treated as definitive. They represent almost a third of all mutation carriers in the cohort.

Expressed as a post-test probability, a negative plasma result moves a patient from a 40% pre-test probability to 21.6%. That is a real reduction and the test is not useless. But a residual probability of one in five is not a basis for withholding a therapy, which is what treating the result as definitive would amount to.

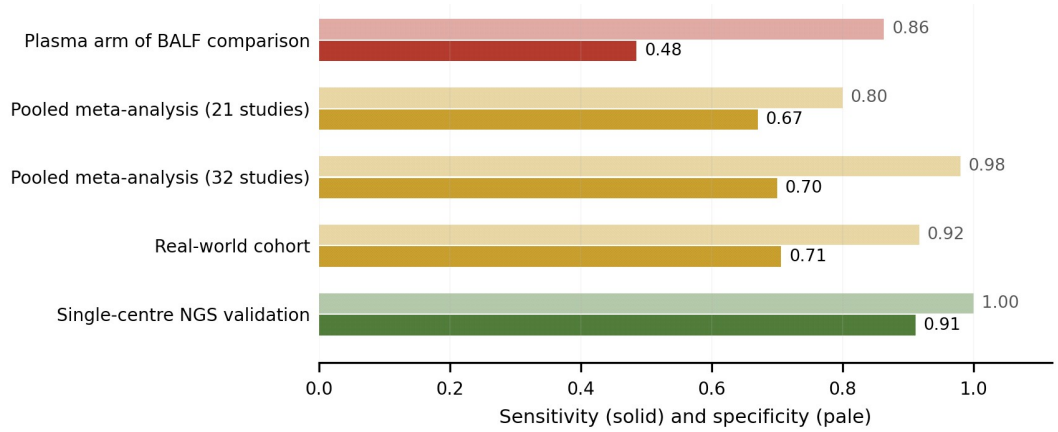
3.5. The pooled figure describes no particular assay

A further difficulty appears when the contributing sources are compared directly, shown in Figure 3. Sensitivity for plasma detection of EGFR alterations against paired tissue ranged from 0.485 in one comparison to 0.911 in a single-centre validation using deep sequencing with a detection limit of 0.01% allele frequency. Specificity ranged from 0.80 to 1.00.

These are not measurements of the same test. They differ in platform, in panel breadth, in detection limit, in disease stage and in the proportion of patients with extrathoracic disease, all of which bear on the amount of tumour DNA in circulation (PMC7262937; Sacher *et al.*, 2016). A pooled sensitivity quoted without those parameters describes an average across assays that a clinician cannot order.

4. Discussion

Plasma genotyping in advanced lung cancer is conventionally summarised by a concordance rate with tissue, and figures between 85% and 97% are widely quoted. The pooled accuracy underlying the most rigorously studied application—detection of the T790M resistance mutation—is a sensitivity of 0.67 and a negative predictive value of 0.60. Both statements describe the same test.



Sensitivity for plasma detection of EGFR alterations against a paired tissue reference ranges from 0.49 to 0.91 across sources. A pooled figure quoted without the assay, platform and detection limit that produced it does not describe any particular test.

Figure 3: Sensitivity (solid bars) and specificity (pale bars) reported across contributing sources for plasma detection of EGFR alterations against paired tissue. Sensitivity ranges from 0.49 to 0.91. These are different assays with different detection limits applied to different populations, and no pooled figure describes any one of them

Table 2: Derived quantities across prevalence, at fixed pooled accuracy

Prevalence	Concordance	Positive predictive value	Negative predictive value
5%	79.3%	15.0%	97.9%
10%	78.7%	27.1%	95.6%
20%	77.4%	45.6%	90.7%
35%	75.4%	64.3%	81.8%
40%	74.8%	69.1%	78.4%
50%	73.5%	77.0%	70.8%
70%	70.9%	88.7%	51.0%
90%	68.3%	96.8%	21.2%

Note: Computed at sensitivity 0.67 and specificity 0.80, held fixed throughout. Concordance varies by 11% points across the entire range while negative predictive value varies by 77. The pooled negative predictive value of 0.60 reported in the source corresponds to a prevalence somewhat above 50%, reflecting the enriched populations in which T790M testing is performed.

The reconciliation is arithmetic. Concordance counts agreement on absence identically to agreement on presence, and in most tested populations the majority of patients do not carry the alteration in question. A test that identified no mutations at all would achieve a concordance equal to one minus the prevalence, which in many series exceeds 80%. Concordance is therefore not a demanding standard, and its stability across populations is a symptom of its insensitivity rather than a sign of robustness.

What clinicians need is conditional and asymmetric. A positive plasma result is highly actionable: the positive predictive value of 0.85 means most positives are real, and the patient can proceed to matched therapy without waiting for tissue. That is a genuine and substantial benefit, and it accounts for much of the enthusiasm for liquid biopsy. A negative result is a different matter. At realistic prevalence it leaves roughly one patient in five still carrying the mutation, and treating it as definitive would deny a superior therapy to nearly a third of eligible patients.

This asymmetry has a clean practical form that most guidance already endorses and that concordance reporting obscures: plasma first, tissue if plasma is negative (Rolfo *et al.*, 2021). Under that algorithm the test earns its place by accelerating treatment for the subset who test positive, without any claim that a negative result excludes anything. The concordance figure is used to argue for something stronger, namely that plasma can substitute for tissue, and the accuracy data do not support that.

The variability between assays compounds the problem. Sensitivity across the sources examined ranged from 0.49 to 0.91, driven by platform, detection limit, panel design and the tumour burden of the population. A single-centre validation reporting 0.91 used deep sequencing capable of detecting variants at 0.01% allele frequency; the pooled figure of 0.67 averages across platforms of widely differing capability. Quoting either to describe a specific clinical assay is unjustified without knowing which.

The reporting recommendation follows directly. Studies of plasma genotyping should report sensitivity, specificity and both predictive values with the prevalence of the tested population stated, together with the platform and limit of detection. Concordance may be reported alongside but should not stand alone, and chance-corrected agreement should accompany it where it is given, since in one contributing cohort a raw concordance of 87.4% corresponded to a kappa of 0.60.

5. Limitations

- The derivations hold sensitivity and specificity fixed while varying prevalence. In practice both vary with disease stage and tumour burden, since circulating tumour DNA fraction depends on them, so the curves are illustrative of the arithmetic rather than predictive of any specific population.
- The pooled accuracy figures used for derivation come principally from one meta-analysis addressing one alteration in one disease. Other alterations, panels and tumour types will differ.
- Derived quantities are deterministic functions of the reported inputs, so no confidence intervals are placed around them. Their uncertainty is that of the pooled sensitivity and specificity.
- The assumed prevalence of 40% for the worked example approximates reported T790M frequency at progression but varies between series and by prior therapy.
- Partial verification bias is a recognised threat: where tissue genotyping is obtained selectively after a negative plasma result, sensitivity estimates are distorted, and the contributing sources did not uniformly report whether tissue was sought in all patients.
- Contributing sources differ in platform, panel breadth and limit of detection, which are the principal determinants of sensitivity and could not be examined systematically from the available data.
- Tissue is treated as the reference standard throughout, but tissue genotyping is itself imperfect and subject to sampling error from intratumoural heterogeneity. Some plasma-positive, tissue-negative results may represent true alterations missed by the biopsy rather than false positives.
- Turnaround time, cost, procedural complication rates and patient acceptability all bear on the choice between plasma and tissue and were outside scope.
- No new pooling was performed, so no heterogeneity or publication bias statistics are reported.
- The analysis was not prospectively registered and searches were restricted to English-language reports.

6. Conclusion

Plasma genotyping against tissue reports concordances of 85% to 97%, and pooled sensitivity of 0.67 with a negative predictive value of 0.60 for the most studied application. These are not in conflict: concordance counts agreement on absence identically to agreement on presence, varies by only 12% points across the entire plausible prevalence range, and would exceed 80% for a test that detected nothing in a population where most patients are mutation-negative.

A positive plasma result is actionable and allows treatment to begin without waiting for tissue. A negative result is not: at realistic prevalence roughly one patient in five with a negative plasma test still carries the alteration, and per 100 patients tested 13 of 40 true carriers are missed. Reflex tissue genotyping after a

negative plasma result therefore remains necessary, and reporting should give sensitivity, predictive values, the prevalence of the tested population and the assay limit of detection rather than a concordance percentage.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

[Author list to be completed from source record]. Extracellular vesicle-based bronchoalveolar lavage fluid liquid biopsy for EGFR mutation testing in advanced non-squamous NSCLC. [PMC9179452].

[Author list to be completed from source record]. High MAF of EGFR mutations and high ratio of T790M sensitizing mutations in ctDNA predict better third-generation TKI outcomes. [PMC7262937].

[Author list to be completed from source record]. (2021). Liquid biopsy for therapy monitoring in early-stage non-small cell lung cancer. *Mol Cancer*. doi: 10.1186/s12943-021-01371-1.

[Author list to be completed from source record]. Real-world data of the correlation between EGFR determination by liquid biopsy in non-squamous non-small cell lung cancer and the EGFR profile in tumor biopsy. [PMID 30847713].

[Author list to be completed from source record]. (2018). The diagnostic accuracy of circulating tumor DNA for the detection of EGFR-T790M mutation in NSCLC: A systematic review and meta-analysis. *Sci Rep*, 8: 13379. doi: 10.1038/s41598-018-30780-4.

[Author list to be completed from source record]. Validation of liquid biopsy: Plasma cell-free DNA testing in clinical management of advanced non-small cell lung cancer. *Lung Cancer (Auckl)*. doi: 10.2147/LCTT.S147841.

Landis, J.R. and Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1): 159-174.

Leeflang, M.M.G., Rutjes, A.W.S., Reitsma, J.B., Hooft, L. and Bossuyt, P.M.M. (2013). Variation of a test's sensitivity and specificity with disease prevalence. *CMAJ*, 185(11): E537-E544.

McInnes, M.D.F., Moher, D., Thombs, B.D., McGrath, T.A. and Bossuyt, P.M. (2018). Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: The PRISMA-DTA statement. *JAMA*, 319(4): 388-396.

Oxnard, G.R., Thress, K.S., Alden, R.S., Lawrance, R., Paweletz, C.P., Cantarini, M. *et al.* (2016). Association between plasma genotyping and outcomes of treatment with osimertinib (AZD9291) in advanced non-small-cell lung cancer. *J Clin Oncol*, 34(28): 3375-3382.

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D. *et al.* (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372: n71.

Rolfo, C., Mack, P., Scagliotti, G.V., Aggarwal, C., Arcila, M.E., Barlesi, F. *et al.* (2021). Liquid biopsy for advanced NSCLC: A consensus statement from the International Association for the Study of Lung Cancer. *J Thorac Oncol*, 16(10): 1647-1662.

Sacher, A.G., Paweletz, C., Dahlberg, S.E., Alden, R.S., O'Connell, A., Feeney, N. *et al.* (2016). Prospective validation of rapid plasma genotyping for the detection of EGFR and KRAS mutations in advanced lung cancer. *JAMA Oncol*, 2(8): 1014-1022.

Whiting, P.F., Rutjes, A.W.S., Westwood, M.E., Mallett, S., Deeks, J.J., Reitsma, J.B. *et al.* (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*, 155(8): 529-536.

Cite this article as: Sarah Williams (2026). Concordance is the wrong metric: Plasma against tissue genotyping in advanced lung cancer, and what a negative liquid biopsy actually rules out. *Biomed Pharma Reviews.* 2(1), 10-16. doi: 10.62587/BMPR.2.1.2026.10-16.