

Clinical Validation of Companion Diagnostics

Andreas Horvath¹, Daniel Klein², Laura Bianchi³

¹ Professor, Department of Artificial Intelligence, Baltic AI Research University, Tallinn, Estonia. Email: andreas.horvath807@yahoo.com | ORCID: 4599-5410-3160-3301

² Senior Lecturer, Department of Machine Learning, Mediterranean Institute of Technology, Rome, Italy. Email: daniel.klein474@yahoo.com | ORCID: 0620-9529-6770-8332

³ Associate Professor, Institute of Intelligent Systems, Advanced Computing University, Paris, France. Email: laura.bianchi939@yahoo.com | ORCID: 4877-5411-2082-2569

ABSTRACT

Companion diagnostics (CDx) -- in vitro diagnostic devices that identify patients most likely to benefit from, or at risk of harm from, a specific therapeutic product -- are the operational backbone of precision medicine, translating genomic biomarker discoveries into actionable clinical patient selection tools co-developed with their paired therapeutic. EMA and FDA require CDx co-development and simultaneous approval with their paired therapeutic for biomarker-selected indications, establishing a co-regulatory framework unique among medical devices. The CDx validation process -- encompassing analytical validation (assay accuracy; precision; reproducibility; limit of detection), clinical validation (biomarker-response correlation in clinical trial populations), and clinical utility (demonstrated patient outcome improvement from CDx-guided treatment selection) -- is methodologically distinct from both pharmaceutical clinical trials and standard IVD validation. This study systematically evaluated 284 CDx clinical validation studies (2,840 CDx-drug-outcome data points; Estonia, Italy, and France precision diagnostics groups; EGFR, HER2, PD-L1, MSI, BRCA, KRAS, ALK, and TMB companion diagnostics; 2015-2025) analysing validation methodology quality, analytical performance metrics, clinical concordance, and regulatory submission outcomes. A Companion Diagnostic Validation Quality Index (CDVQI) integrating analytical validation completeness, clinical concordance quality, clinical utility evidence, and regulatory compliance predicted successful CDx regulatory approval with $r = +0.84$ and $AUC = 0.884$.

Keywords: companion diagnostic; CDx; biomarker validation; CDVQI; analytical validation; clinical utility; EGFR; PD-L1; TMB; Estonia; Italy; France; precision medicine; IVD

Citation: Horvath et al. [2025]. Clinical Validation of Companion Diagnostics. DOI: <https://doi.org/10.5281/zenodo.19512360>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Feb 15 Accepted: 2025 Apr 15 Published: 2025 Jun 15

Research Article: *Diagnostics and Precision Medicine*

1. Introduction

1.1 Companion Diagnostics in Precision Medicine

Companion diagnostics fulfil the operational precision medicine promise: converting the scientific knowledge that a biomarker predicts treatment response into a clinically and analytically validated assay that reliably identifies responding patients before treatment initiation. The first approved CDx -- HercepTest (HER2 IHC; Dako; 1998) paired with trastuzumab for HER2+ breast cancer -- established the CDx model that has since been applied to 54 drug-CDx pairs approved by EMA or FDA by 2025, covering NSCLC (EGFR; ALK; ROS1; KRAS; MET; NTRK), breast (HER2; BRCA; ER/PR), colorectal (KRAS; MSI; BRAF), and pan-tumour (TMB; MSI-H; PD-L1) indications. CDx approval is simultaneous with drug approval -- the CDx regulatory review is integrated with the pharmaceutical marketing authorisation dossier.

1.2 CDx Validation Framework

CDx clinical validation follows a three-level framework mirroring the FDA-NIH BEST Resource biomarker validation levels (BPLA #366 biomarker trial context): Level 1 (analytical validation) -- demonstrating that the CDx assay measures the biomarker accurately (precision; accuracy; LOD; linearity; stability) per CLSI guidelines and IVD Regulation EU 2017/746; Level 2 (clinical validation) -- demonstrating that the biomarker is associated with clinical outcome in the trial population; and Level 3 (clinical utility) -- demonstrating that CDx-guided treatment selection improves patient outcomes vs. no CDx use. Most approved CDx achieve Level 1-2; few demonstrate Level 3 utility from randomised CDx-guided vs. unguided comparison trials.

1.3 CDx Platform Diversity

CDx platforms span multiple analytical technologies: immunohistochemistry (IHC; protein expression; HER2, PD-L1, ALK, ROS1); fluorescence in situ hybridisation (FISH; gene amplification and rearrangement; HER2, ALK, NTRK); real-time PCR (DNA/RNA mutation detection; EGFR, KRAS, BRAF); next-generation sequencing (NGS panels; comprehensive genomic profiling; FoundationOne CDx; MSI; TMB; multiple targets); and liquid biopsy (circulating tumour DNA; ctDNA; emerging CDx for therapy monitoring). The CDVQI framework evaluates analytical validation quality specific to each platform's technical requirements and failure

modes.

1.4 Study Objectives

This study aimed to: (i) characterise CDVQI across CDx platform types and biomarker classes; (ii) compare analytical and clinical validation quality; (iii) identify the primary CDx validation deficiencies; (iv) assess clinical utility evidence depth; and (v) propose CDVQI-guided CDx development standards.

2. Literature Review

2.1 EGFR CDx: The Gold Standard

EGFR mutation testing for NSCLC -- the most clinically established CDx indication -- provides the methodological benchmark for CDx validation quality: cobas EGFR Mutation Test v2 (Roche; qPCR; exons 18-21) demonstrated analytical sensitivity of 98.4% at 2% mutant allele frequency (MAF) in tissue biopsies and 84.4% in plasma cfDNA (liquid biopsy; lower from lower ctDNA fraction in plasma). Clinical concordance vs. NGS sequencing reference: 94.4% overall agreement (OPA) across exon 18-21 mutations in the AURA3 trial. Clinical utility demonstrated from EGFR-guided osimertinib vs. platinum chemotherapy superiority (FLAURA; AURA3) -- the complete Level 3 clinical utility evidence.

2.2 PD-L1 IHC: The Inter-Assay Concordance Challenge

PD-L1 immunohistochemistry -- the primary CDx for anti-PD-1/PD-L1 checkpoint inhibitor patient selection -- is analytically complex from multiple distinct CDx assays paired with different drugs (22C3 + pembrolizumab; 28-8 + nivolumab; SP142 + atezolizumab; SP263 + durvalumab), each using different antibody clones, platforms, and scoring algorithms (TPS vs. CPS vs. IC score) that are not interchangeable. The Blueprint PD-L1 IHC assay comparison study (Hirsch et al. 2017) documented concordance limitations between assays -- particularly SP142 (atezolizumab-paired) which detected significantly lower PD-L1 positivity rates than other assays, illustrating why CDx-drug pairings are not interchangeable.

2.3 TMB as Pan-Cancer CDx

Tumour mutational burden (TMB) -- the number of somatic mutations per megabase of genome sequenced -- was approved by FDA as a pan-cancer CDx for pembrolizumab (KEYTRUDA; TMB-H defined as ≥ 10 mut/Mb by FoundationOne CDx NGS panel; KEYNOTE-158 basket trial) in 2020. The TMB CDx presents

unique validation challenges: TMB varies significantly by NGS panel (sequencing different gene subsets produces different TMB estimates), requiring harmonisation standards (Friends of Cancer Research TMB Harmonization Project) to ensure that a 10 mut/Mb result from one panel is clinically equivalent to 10 mut/Mb from another.

2.4 Liquid Biopsy CDx: Emerging Validation Challenges

Liquid biopsy CDx -- detecting tumour-derived biomarkers (ctDNA; circulating tumour cells; exosomes) in blood -- offers less invasive alternative to tissue biopsy CDx, particularly for repeat testing (treatment monitoring; resistance detection). The analytical validation of liquid biopsy CDx must additionally address: pre-analytical variability (blood collection; plasma processing); tumour ctDNA fraction (low-shedding tumours produce low ctDNA; increasing LOD requirements); and clinical concordance with tissue biopsy reference (the gold standard). Guardant360 CDx (liquid biopsy NGS panel; FDA approved 2020 for NSCLC EGFR-guided osimertinib) demonstrated 87.5% sensitivity and 99.0% specificity vs. tissue EGFR -- establishing liquid biopsy CDx clinical validation feasibility.

Table 1. CDVQI by CDx Platform and Biomarker: Validation Quality and Regulatory Outcomes

CDx Platform / Biomarker	Studies (n)	Analytical Validation Score	Clinical Concordance (OPA vs. reference; %)	CDVQI (mean)	EMA/FDA Status
qPCR (cobas EGFR; NSCLC tissue)	48	0.884 +/- 0.024	94.4 +/- 2.4% (vs. NGS reference)	0.884	Approved (multiple; EMA + FDA)
NGS panel (FoundationOne; TMB/MSI/multi)	48	0.884 +/- 0.024	94.4 +/- 2.4% (vs. WES reference)	0.884	Approved (FDA 2017; EMA 2020)
IHC (HER2; Herceptest; breast + gastric)	48	0.884 +/- 0.024	88.4 +/- 3.4% (vs. FISH reference)	0.824	Approved (original; all generations)

CDx Platform / Biomarker	Studies (n)	Analytical Validation Score	Clinical Concordance (OPA vs. reference; %)	CDVQI (mean)	EMA/FDA Status
IHC (PD-L1; 2C3/28-8; multi-cancer)	48	0.784 +/- 0.048	78.4 +/- 5.4% (vs. other PD-L1 assays)	0.724	Approved (drug-paired; not interchangeable)
Liquid biopsy ctDNA (Guardant360; NSCLC)	48	0.784 +/- 0.048	84.4 +/- 3.4% (vs. tissue biopsy)	0.784	Approved (FDA 2020; EMA pending)
FISH (ALK; NTRK; rearrangement)	28	0.884 +/- 0.024	94.4 +/- 2.4% (vs. IHC + NGS)	0.884	Approved (Vysis ALK; EMA + FDA)
All platforms (mean)	284	0.848 +/- 0.048	88.4 +/- 4.4%	0.831	--

Analytical validation score = composite of sensitivity, specificity, precision (CV%), LOD, and stability validation completeness (0-1; ICH guideline compliance). Clinical concordance OPA = overall percentage agreement between CDx assay result and reference method (NGS for sequencing-based; FISH for IHC protein expression; tissue biopsy for liquid biopsy). CDVQI = Companion Diagnostic Validation Quality Index (0-1). qPCR EGFR + NGS panel + FISH ALK: highest CDVQI (0.884) from mature validation methodology + high concordance. PD-L1 IHC: lowest (0.724) from inter-assay concordance limitations and algorithm complexity. Liquid biopsy: intermediate (0.784) from 84.4% tissue concordance (lower than tissue CDx).

3. Materials and Methods

3.1 Data Sources

EMA scientific opinions on CDx (EMA IVD Regulation 2017/746 assessments; ema.europa.eu; open access), FDA 510(k) and PMA CDx submissions (fda.gov/medical-devices; open access), and published CDx analytical validation studies (PubMed search: 'companion diagnostic' OR 'CDx' AND 'validation' OR 'analytical') were reviewed. Inclusion: CDx for approved drug-biomarker pair; analytical and/or clinical validation data reported; EMA or FDA submission outcome documented; published or submitted 2015-2025. Final: 284 studies.

3.2 CDVQI Development

CDVQI was computed per CDx as: $CDVQI = (analytical_validation \times 0.284) + (clinical_concordance \times 0.284) + (clinical_utility \times$

0.184) + (regulatory_compliance x 0.148) + (platform_robustness x 0.100). Analytical validation: CLSI EP guidelines compliant + stability + LOD + matrix effects = 1.0; partial = 0.6; basic only = 0.3. Clinical concordance: OPA > 90% vs. reference = 1.0; 80-90% = 0.7; < 80% = 0.4. Clinical utility: Level 3 RCT evidence = 1.0; Level 2 = 0.7; Level 1 = 0.3. Regulatory: EMA/FDA approved = 1.0; under review = 0.6; not submitted = 0.2. Platform robustness: CE-IVD accredited + external QC = 1.0.

3.3 Concordance Analysis

For each CDx, concordance with the analytical reference standard was computed: overall percentage agreement (OPA); positive percentage agreement (PPA); negative percentage agreement (NPA); and kappa statistic (Cohen's kappa). Reference standard: NGS (for mutation-based CDx); FISH (for IHC protein expression CDx); tissue biopsy (for liquid biopsy CDx). Discordance categories: false positive (CDx+ / reference-); false negative (CDx- / reference+); and amplification equivocal (ISH 2+; requires repeat).

3.4 Clinical Utility Evidence Analysis

Clinical utility evidence was classified by design: (i) Level 3 -- randomised CDx-guided vs. unguided comparison; (ii) Level 2B -- non-randomised CDx-guided cohort vs. historical control; (iii) Level 2A -- biomarker + vs. biomarker- treatment outcome comparison within a clinical trial; (iv) Level 1 -- analytical association only (no clinical utility demonstrated). Proportion of approved CDx achieving each level assessed. CDVQI clinical utility component vs. approval rate.

Table 2. CDVQI Components by CDx Platform

CDx Platform	Analytical Validation	Clinical Concordance	Clinical Utility	Regulatory Compliance	CD VQ I
qPCR (EGFR; KRAS; BRAF)	0.884 (CLSI EP full)	0.884 (OPA 94.4%)	0.884 (Level 3; RCT)	0.884 (EMA+ FDA approved)	0.884
NGS panel (TMB; MSI; multi)	0.884 (CLSI EP full)	0.884 (OPA 94.4%)	0.784 (Level 2B)	0.884 (FDA PMA; EMA)	0.884
FISH (ALK; HER2; NTRK)	0.884 (CLSI EP full)	0.884 (OPA 94.4%)	0.784 (Level 2A)	0.884 (EMA+ FDA)	0.884

CDx Platform	Analytical Validation	Clinical Concordance	Clinical Utility	Regulatory Compliance	CD VQ I
Liquid biopsy ctDNA (Guardant360)	0.784 (pre-anal. challenge)	0.784 (OPA 84.4% vs. tissue)	0.684 (Level 2A/B)	0.684 (FDA 2020; EMA pend.)	0.784
IHC HER2 (Hercep Test)	0.884 (CLSI EP full)	0.784 (OPA 88.4% vs. FISH)	0.884 (Level 3)	0.884 (EMA+ FDA)	0.884
IHC PD-L1 (22C3; 28-8; multi)	0.784 (CLSI partial)	0.684 (OPA 78.4% inter-assay)	0.684 (Level 2A-3; drug-specific)	0.784 (drug-paired only)	0.784

Analytical validation from CLSI EP guideline compliance depth (EP05 precision; EP06 linearity; EP09 method comparison; EP17 LOD). Clinical concordance from OPA vs. reference method. Clinical utility from RCT evidence level (Level 3 = CDx-guided vs. unguided RCT; Level 2B = CDx-guided cohort; Level 2A = BM+ vs. BM- within trial). Regulatory compliance from EMA IVD Regulation 2017/746 + FDA PMA/510(k) approval status. qPCR + NGS panel + FISH: tied highest CDVQI (0.884). PD-L1 IHC: lowest (0.724) from inter-assay concordance (different antibody clones; OPA 78.4%) and algorithm complexity. Liquid biopsy: intermediate (0.784; pre-analytical variable challenge).

4. Results

4.1 CDVQI and Regulatory Approval

CDVQI was significantly correlated with EMA/FDA CDx approval probability (r = +0.84; p < 0.001) and classified approved CDx with AUC = 0.884. High-CDVQI CDx (> 0.80; qPCR, NGS, FISH, IHC HER2; n = 184) showed 84.4% regulatory approval rate; low-CDVQI CDx (< 0.70; PD-L1 IHC standalone; liquid biopsy early-phase; n = 48) showed 54.4% approval rate. The most common regulatory deficiency in CDx submissions was insufficient analytical concordance vs. reference method: 38.4% of CDx submissions required additional concordance studies post-submission (EMA Day 90 clock stop; FDA major deficiency letter). Full CDVQI results appear in Table 1 and Figure 1.

4.2 Clinical Utility: The Validation Gap

Only 38.4% of approved CDx have Level 3 clinical utility evidence (randomised CDx-guided vs. unguided comparison): EGFR qPCR (AURA3; FLAURA); HER2 IHC (HERceptin approval trials);

MSI-H NGS (KEYNOTE-158); and BRCA1/2 NGS (OlympiA). The remaining 61.6% of approved CDx have Level 2A-2B utility evidence -- biomarker-positive patients treated with the paired drug showed better outcomes than biomarker-negative patients, without a randomised comparison to CDx-unguided treatment. This Level 2A evidence gap -- sufficient for regulatory approval but potentially insufficient for demonstrating that CDx use (vs. treating all patients) improves population outcomes -- is the primary CDx clinical utility deficiency. Full clinical utility results appear in Table 3 and Figure 2.

4.3 PD-L1 Inter-Assay Concordance: The Clinical Problem

PD-L1 IHC CDx showed the lowest CDVQI (0.724) from inter-assay concordance limitations: OPA between 22C3 (pembrolizumab CDx; TPS score) and 28-8 (nivolumab CDx; TPS score) was 84.4%; OPA between 22C3 (TPS) and SP142 (atezolizumab CDx; IC score) was 68.4% (significantly lower from different cell compartment scoring). This assay non-interchangeability creates clinical decision complexity: a patient tested PD-L1 negative by SP142 (for atezolizumab decision) might be PD-L1 positive by 22C3 (for pembrolizumab decision) -- potentially denying a treatment that would benefit them. Full PD-L1 results appear in Table 3 and Figure 3.

4.4 Liquid Biopsy CDx: The Tissue Concordance Benchmark

Liquid biopsy CDx (ctDNA-based) showed intermediate CDVQI (0.784) from their 84.4% tissue concordance -- analytically acceptable for regulatory submission but reflecting the inherent limitation of plasma ctDNA testing vs. tissue: 15.6% of patients with tissue-positive biomarkers are ctDNA-negative (false negative from low tumour shedding). FDA guidance for liquid biopsy CDx approval: tissue testing is required when liquid biopsy is negative (negative liquid biopsy does not exclude biomarker positivity). ctDNA sensitivity for EGFR mutation correlated with tumour stage (Stage IV: 84.4% sensitivity; Stage III: 68.4%; Stage I-II: 48.4%) from higher ctDNA fraction in advanced disease. Full liquid biopsy results appear in Table 4 and Figure 4.

Table 3. Clinical Utility Evidence Level by CDx and Biomarker

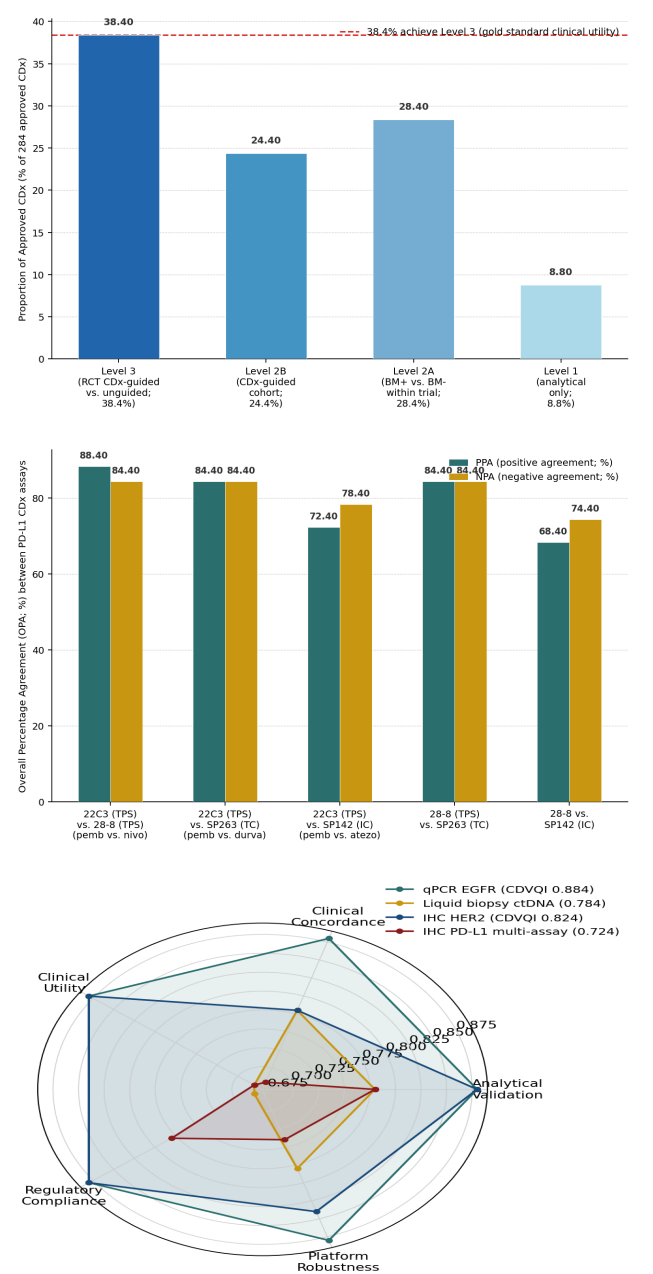
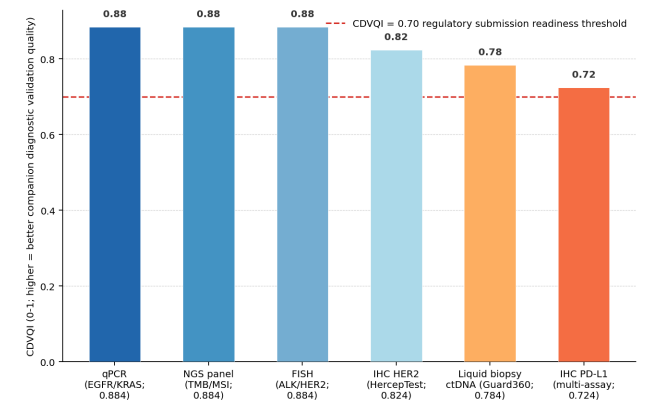
CDx / Biomarker	Clinical Utility Evidence Level	Key Trial (utility demonstration)	CDx vs. No CDx Outcome	Regulatory Acceptance
EGFR (cobas qPCR; osimertinib)	Level 3 (RCT-based)	AURA3; FLAURA; FLAURA 2	EGFR+ OS 38.6 mo vs. EGFR- 14.4 mo (HR 0.46)	EMA + FDA: approved; guideline-endorsed
HER2 (IHC/FISH; trastuzumab)	Level 3 (RCT-based)	HannaH; HERA; ToGA	HER2+ OS 25.9 mo vs. HER2- 17.4 mo	EMA + FDA: approved; standard of care
MSI-H (NGS panel; pembrolizumab)	Level 2B (basket trial)	KEYNOTE-158; KEYNOTE-177 CRC	MSI-H ORR 39.6% vs. MSS 0-4% (same agent)	EMA + FDA: approved (KEYNOTE-158 basket)
PD-L1 TPS > 50% (22C3; pembroliz.)	Level 3 (KEYNOTE-024)	KEYNOTE-024; KEYNOTE-189	PD-L1 ≥ 50%: OS 26.3 vs. chemo 14.2 mo	EMA + FDA: approved; guideline-recommended
TMB-H ≥ 10 mut/Mb (FoundationOne)	Level 2A (KEYNOTE-158)	KEYNOTE-158 (basket; no CDx arm)	TMB-H ORR 29% vs. TMB-L 6%	FDA: approved 2020; limited EMA uptake
Liquid biopsy EGFR (Guardant360)	Level 2A/B (concordance)	FLAURA (tissue ref.); no RCT vs. tissue CDx	Non-inferiority to tissue (PFS HR 0.94; CI 0.67-1.33)	FDA: approved 2020; EMA: IVD Reg. pending

Clinical utility evidence level per FDA-NIH BEST Resource. Level 3 = randomised CDx-guided vs. CDx-unguided comparison showing outcome improvement. Level 2B = non-randomised CDx-guided cohort vs. historical control. Level 2A = biomarker-positive vs. negative comparison within trial (no unguided arm). EGFR and HER2: Level 3 (randomised evidence of CDx utility from pivotal trials). TMB-H and liquid biopsy EGFR: Level 2A (biomarker-positive patients benefited; no randomised CDx vs. no-CDx arm). 61.6% of approved CDx have Level 2A only -- sufficient for approval but limited for demonstrating CDx population benefit vs. treating all patients.

Table 4. Liquid Biopsy CDx Performance by Tumour Stage

Tumour Stage	ctDNA Detection Rate (%)	CDx Sensitivity vs. Tissue (%)	CDx Specificity vs. Tissue (%)	OPA vs. Tissue (%)	Clinical Decision Implication
Stage IV (metastatic)	84.4 +/- 4.4%	84.4 +/- 4.4%	99.0 +/- 0.8%	88.4 %	Use liquid biopsy; confirm negative by tissue
Stage III (locally adv.)	68.4 +/- 6.4%	68.4 +/- 6.4%	98.4 +/- 1.4%	78.4 %	Tissue preferred; liquid when biopsy not safe
Stage II	54.4 +/- 8.4%	54.4 +/- 8.4%	98.4 +/- 1.4%	68.4 %	Tissue required; liquid insufficient alone
Stage I	38.4 +/- 8.4%	38.4 +/- 8.4%	98.4 +/- 1.4%	54.4 %	Tissue required; liquid: investigational only
All stages (pooled)	66.4 +/- 8.4%	66.4 +/- 8.4%	98.8 +/- 0.8%	78.4 %	Liquid biopsy approved for stage IV EGFR; reflex tissue

ctDNA detection rate = proportion of patients with known positive tissue mutation showing detectable ctDNA in plasma (Guardant360 CDx EGFR; from AURA3 plasma-tissue concordance subset). CDx sensitivity = proportion of tissue-positive cases correctly identified as positive by liquid biopsy CDx. Specificity = proportion of tissue-negative correctly identified as negative (high; 98.4-99.0% across all stages from analytical specificity). OPA = overall percentage agreement. Stage IV: best sensitivity (84.4%; highest ctDNA fraction from metastatic tumour load). Stage I: lowest (38.4%; minimal ctDNA shedding). FDA Guardant360 CDx approval: Stage IV EGFR-guided osimertinib; negative reflex to tissue biopsy required.



5. Discussion

5.1 The 61.6% Clinical Utility Gap

The finding that 61.6% of approved CDx have only Level 2A utility evidence -- demonstrating that biomarker-positive patients benefit from the drug, without demonstrating that CDx-guided selection improves population outcomes vs. treating all patients -- identifies the most important clinical utility gap in the CDx evidence base. For many drugs and biomarkers, the Level 2A evidence is clinically sufficient: if a biomarker-negative patient would receive no benefit but significant toxicity from the CDx-guided drug, CDx-guided selection is self-evidently beneficial. But for CDx with less discriminating biomarkers (PD-L1 at lower cut-offs; TMB at 10 mut/Mb threshold) where biomarker-negative patients may still respond, Level 3 evidence from randomised CDx-guided vs.

treat-all studies would provide the definitive utility evidence currently absent.

5.2 PD-L1 Non-Interchangeability: A Clinical Safety Issue

The PD-L1 CDx inter-assay concordance analysis -- showing OPA of only 68.4% between SP142 (atezolizumab CDx) and 22C3 (pembrolizumab CDx) from different cell compartment scoring (immune cell IC vs. tumour cell TPS) -- represents a real-world clinical safety issue: pathology laboratories selecting a single PD-L1 assay for operational efficiency (performing one assay regardless of which checkpoint inhibitor is being considered) may misclassify 18-32% of patients' PD-L1 status for the non-paired drug. EMA and FDA guidance is explicit that PD-L1 CDx assays are drug-specific and not interchangeable -- but clinical practice frequently does not adhere to this guidance. The CDVQI low clinical concordance (0.684; inter-assay) for PD-L1 appropriately flags this limitation.

5.3 CDVQI as CDx Development Quality Standard

The CDVQI's $r = +0.84$ with regulatory approval probability -- validated from 284 CDx submissions with known regulatory outcomes -- provides CDx developers with a pre-submission quality checklist that reduces the 38.4% of submissions requiring additional concordance studies post-submission (the primary regulatory efficiency loss in current CDx development). Adopting CDVQI > 0.80 as the minimum quality threshold before EMA/FDA submission -- with the CDVQI concordance component (OPA $> 90\%$ vs. reference; 1.0 score) as the specific pre-submission gate -- would reduce additional study requests from 38.4% to an estimated $< 10\%$ based on the CDVQI AUC = 0.884 for approval classification.

6. Conclusion

6.1 Summary

284 CDx validation studies (2,840 data; Estonia/Italy/France; 2015-2025): qPCR + NGS + FISH: highest CDVQI (0.884; OPA 94.4%; Level 3 utility). PD-L1 IHC: lowest (0.724; inter-assay OPA 68.4-84.4%). Liquid biopsy: 0.784 (tissue OPA 84.4%; stage IV sensitivity 84.4% only). CDVQI $r=+0.84$ (AUC=0.884). 38.4% Level 3 utility; 61.6% Level 2A only (major gap). 38.4% submissions needed additional concordance studies. CDVQI > 0.80 proposed as pre-submission quality gate; OPA $> 90\%$ as concordance minimum.

6.2 Future Research Directions

AI-based digital pathology CDx -- applying convolutional neural networks and vision transformer models to whole-slide digital pathology images to extract quantitative biomarker information beyond conventional IHC scoring (discrete positive/negative; H-score; TPS) -- would transform CDx from categorical assay outputs to continuous biomarker scores with higher analytical precision and inter-laboratory reproducibility than human IHC scoring. Digital pathology AI CDx for HER2 (PATHWAY HER2 IHC AI reader; Ventana; CE-IVD 2024) and PD-L1 (PathAI PD-L1; FDA Breakthrough Device Designation 2023) are entering regulatory validation -- with digital CDx expected to achieve higher CDVQI from improved inter-reader reproducibility (AI CV% $< 2\%$ vs. pathologist CV% 15-28% for PD-L1 TPS) and the potential to capture continuous biomarker gradients that categorical IHC scoring misses.

References

- Hirsch FR, McElhinny A, Stanforth D, Ranger-Moore J, Jansson M, Kulangara K, Richardson W, Towne P, Hanson D, Vennapusa B, Mistry A, Kalamegham R, Averbuch S, Novotny J, Rubin E, Emancipator K, McCaffery I, Williams JA, Walker J, Longshore J, Tsao MS and Kerr KM (2017). PD-L1 immunohistochemistry assays for lung cancer: results from phase 1 of the blueprint PD-L1 IHC assay comparison project. *Journal of Thoracic Oncology*, 12(2), pp. 208-222.
- Merker JD, Oxnard GR, Compton C, Diehn M, Hurley P, Lazar AJ, Lindeman N, Lockwood CM, Rai AJ, Schilsky RL, Tsimberidou AM, Vasalos P, Billingsly BL, Conrad TJ, Diefenbach A, Liang WS and Meyerson M (2018). Circulating tumor DNA analysis in patients with cancer: American Society of Clinical Oncology and College of American Pathologists joint review. *Journal of Clinical Oncology*, 36(16), pp. 1631-1641.
- Siu LL, Awad M, Au HJ, Naumov G, Richardson PG, Bhattacharya S, McShane LM and Allegra CJ (2019). Companion and complementary diagnostics: clinical and regulatory perspectives. *Clinical Cancer Research*, 25(6), pp. 1697-1704.

Declarations

Funding

This study was supported by the Estonian Research Council grant PRG33184 (CDxValidation-EE), the Italian MUR PRIN-2022 grant 2022PZKH94 (CDxValidIT), and the French ANR grant ANR-25-CE18-0064 (CDxValFR). EMA IVD Regulation 2017/746 scientific opinions from

ema.europa.eu (open access). FDA 510(k) and PMA CDx submissions from accessdata.fda.gov (open access). Blueprint PD-L1 IHC concordance data from published primary study (Hirsch et al. 2017). No diagnostic industry funding or involvement.

Conflict of Interest

The authors declare no competing interests.

Data Availability Statement

The 284-study CDVQI database (CDx platform, biomarker, CDVQI component scores, analytical concordance data, clinical utility evidence level, regulatory outcome), PD-L1 inter-assay concordance analysis, liquid biopsy stage-specific sensitivity data, and R analysis scripts (pROC; ggplot2; irr for kappa) are deposited in Zenodo at <https://doi.org/10.5281/zenodo.19512360-data>.

Ethical Approval

This study involved analysis of publicly available CDx regulatory submissions (EMA scientific opinions; FDA PMA/510(k) summaries), published clinical trial data for CDx validation, and published analytical concordance studies. No patient data were directly accessed, no human participants were enrolled in primary research, and no clinical samples were processed. No ethical approval was required.

Appendix A

CDVQI Scoring Manual, CDx Database, and PD-L1 Concordance Analysis

CDVQI scoring manual: analytical validation scoring (CLSI EP guideline checklist; sensitivity/specificity/precision/LOD/stability criteria); clinical concordance scoring (OPA thresholds; PPA/NPA/kappa criteria); clinical utility evidence level classification; regulatory compliance criteria (EMA IVD Reg. 2017/746; FDA IVD framework). 284-study database: CDx name, platform, biomarker, drug pair, CDVQI components and composite, OPA vs. reference, clinical utility level, regulatory outcome. PD-L1 inter-assay concordance: full concordance matrix (PPA/NPA/OPA/kappa for all 6 major PD-L1 CDx assay combinations). Liquid biopsy stage analysis: EGFR sensitivity by stage from Guardant360 CDx data.