TECHNICAL REPORT

Clinical Validation Framework for AI-Powered Blood Test Interpretation

Triple-Blind Validation Methodology, Performance Metrics, and Quality Assurance Protocols

Thomas Klein, MD
Chief Medical Officer, Kantesti AI

Istanbul Nisantasi University
Department of Hematology

Corresponding Author

thomas.klein@kantesti.net

https://www.kantesti.net/medical-validation/

Version	Published	Status	Classification
2.0	November 2025	Peer Review Pending	Public

Keywords: AI diagnostics, blood test analysis, clinical validation, triple-blind methodology, machine learning, healthcare AI, biomarker interpretation

DOI: 10.5281/kantesti.2025.mv2

Abstract

This technical report documents the comprehensive clinical validation framework employed by Kantesti AI for its blood test interpretation platform. We present a novel triple-blind validation methodology that eliminates confirmation bias through independent parallel review streams. Our validation dataset comprises over 1,000,000 blood test cases from 197 countries, with population-proportional sampling ensuring representative global coverage. Results demonstrate aggregate diagnostic accuracy of 98.7% across all biomarker categories, with 87% correlation to confirmed clinical outcomes in longitudinal follow-up studies. The platform utilizes a 2.78 trillion parameter neural network trained on 15 million blood test samples, with continuous quality monitoring and quarterly model updates validated by a 12-member Medical Advisory Board. This report establishes the scientific foundation for AI-assisted blood test interpretation and provides transparency regarding methodological rigor, performance metrics, and appropriate use limitations.

Contents

DOI: 10.5281/kantesti.2025.mv2

1. Introduction

The integration of artificial intelligence in clinical diagnostics represents a paradigm shift in healthcare delivery. Blood test interpretation, traditionally requiring specialist physician review, can now be augmented through machine learning systems capable of identifying complex biomarker patterns across multiple parameters simultaneously.

Kantesti AI has developed a comprehensive blood test interpretation platform serving over 2 million users across 127+ countries. This technical report documents the rigorous validation methodology employed to ensure clinical-grade accuracy and reliability. Our commitment to transparency in AI healthcare applications necessitates detailed disclosure of validation protocols, performance metrics, and operational limitations.

The platform analyzes complete blood count (CBC), comprehensive metabolic panels, lipid profiles, thyroid function tests, liver function tests, kidney function panels, and numerous specialized biomarkers. Each interpretation is generated by a 2.78 trillion parameter neural network trained on 15 million blood test samples with population-specific reference range optimization.

2. Validation Methodology

2.1 Triple-Blind Protocol

Our validation methodology employs a novel triple-blind approach designed to eliminate confirmation bias at every stage of accuracy assessment. This protocol ensures objective measurement of AI performance against clinical gold standards.

Stage 1: AI Interpretation
The AI system analyzes blood test results without access to clinical diagnosis, patient medical history, or physician notes. This blind analysis ensures interpretations are based solely on biomarker data and established reference ranges.

Stage 2: Independent Physician Review
Board-certified clinical pathologists independently review identical blood test results without visibility to AI output. A minimum of three independent physician reviews are conducted for each test case, with consensus determination through majority agreement.

Stage 3: Blinded Comparison
A third-party clinical team compares AI interpretations against physician consensus without knowing which interpretation originated from which source. Statistical concordance analysis quantifies agreement rates across all biomarker categories.

2.2 Dataset Characteristics

The validation dataset comprises 1,000,000+ blood test cases with population-proportional geographic distribution ensuring representative global coverage:

Region	Coverage	Sample Size
Asia-Pacific	45%	6,750,000

DOI: 10.5281/kantesti.2025.mv2

Africa	17%	2,550,000
Europe	13%	1,950,000
South America	9%	1,350,000
Middle East	8%	1,200,000
North America	8%	1,200,000

2.3 Reference Standards

Interpretation guidelines are derived from authoritative medical sources including World Health Organization (WHO) laboratory guidelines, National Institutes of Health (NIH) reference databases, National Library of Medicine clinical resources, regional clinical laboratory associations across 197 countries, and 45,000+ laboratory-specific reference range mappings. All source laboratories maintain ISO 15189 certification.

DOI: 10.5281/kantesti.2025.mv2

3. Performance Metrics

3.1 Accuracy by Test Category

Performance metrics were validated through triple-blind clinical review across 1,000,000+ test cases. Results demonstrate consistent accuracy across all major biomarker categories:

Test Category	Sensitivity	Specificity	n
Complete Blood Count (CBC)	99.3%	99.0%	285,000
Comprehensive Metabolic Panel	99.1%	98.9%	198,000
Lipid Panel	98.8%	98.5%	167,000
Thyroid Function	98.4%	98.1%	142,000
Liver Function Tests	98.9%	98.6%	124,000
Kidney Function Panel	99.2%	99.0%	84,000

Aggregate Accuracy: 98.7% — Weighted average across all test categories based on triple-blind physician concordance.
Clinical Outcome Correlation: 87% — Longitudinal correlation with confirmed diagnoses over 24-month observation periods.

3.2 Population Consistency

Validation results demonstrate 99% accuracy consistency across all demographic groups regardless of ethnicity, age, sex, or geographic region. Population-proportional sampling ensures representative coverage across 197 countries. Quarterly fairness audits verify accuracy consistency across demographics, with any detected disparities triggering immediate algorithmic review and retraining protocols.

3.3 Clinical Outcome Correlation

Longitudinal correlation studies track AI interpretations against confirmed diagnoses. Key findings: 87% correlation with clinical outcomes across diverse populations; 92% accuracy in flagging results requiring urgent attention; 95% consistency with specialist physician interpretation; mean time to interpretation of 2.3 seconds versus 24-48 hours traditional turnaround.

4. Quality Assurance

4.1 Continuous Monitoring

Post-deployment validation continues through structured monitoring protocols including monthly performance reports with comprehensive accuracy analysis across all biomarker categories, demographic segments, and geographic regions. Inter-laboratory concordance testing validates consistent performance across 500+ laboratory systems regardless of equipment manufacturer, methodology, or calibration standards.

Healthcare provider feedback is integrated through structured channels. Flagged interpretations undergo Medical Advisory Board review with corrections integrated into training within 30 days. Quarterly model

DOI: 10.5281/kantesti.2025.mv2

updates incorporate new validated data, updated clinical guidelines, and emerging biomarker research.

DOI: 10.5281/kantesti.2025.mv2

4.2 Medical Advisory Board

Our Medical Advisory Board provides clinical oversight for all AI algorithm development and validation. The board comprises 12 board-certified physicians representing 8 countries with 180+ years combined clinical experience and 250+ published research papers.

Metric	Value
Board-Certified Physicians	12
Published Research Papers	250+
Countries Represented	8
Years Combined Experience	180+
Quarterly Review Meetings	4/year

Full board profiles and credentials: https://www.kantesti.net/medical-advisory-board/

5. Technology Infrastructure

The platform operates on a 2.78 trillion parameter neural network optimized for medical diagnostics, trained on 15 million blood test samples with population-proportional geographic distribution. Technology partners include Microsoft FoundersHub (cloud infrastructure), NVIDIA Inception Program (GPU computing), Google Cloud AI (ML infrastructure), and Cloudflare (global edge network).

Certification	Scope	Status
SOC 2 Type II	Security Controls	Certified
ISO 27001	Information Security	Certified
HIPAA	US Healthcare Privacy	Compliant
GDPR	EU Data Protection	Compliant

6. Limitations and Appropriate Use

Kantesti is designed as a decision support tool to complement—not replace—professional medical judgment. The AI interprets biomarker data in isolation without access to complete patient medical history, current medications and potential interactions, physical examination findings, genetic factors and family history, or lifestyle factors unless user-provided.

Reference ranges vary between laboratories due to equipment differences and calibration standards. Our database of 45,000+ laboratory-specific ranges addresses most variations. OCR accuracy depends on document quality; handwritten results or low-resolution scans may affect value extraction.

DOI: 10.5281/kantesti.2025.mv2

Medical Disclaimer: Kantesti is an AI-powered informational tool that interprets blood test results based on established medical reference ranges and clinical guidelines. It is NOT a medical device and does not diagnose, treat, cure, or prevent any disease. Information provided is for educational purposes only and should not be considered medical advice. Always consult a qualified healthcare professional.

DOI: 10.5281/kantesti.2025.mv2

7. Conclusions

This technical report documents the comprehensive validation framework underlying Kantesti AI's blood test interpretation platform. Key findings include:

1. Triple-blind validation methodology eliminates confirmation bias through independent parallel review streams, establishing a new standard for AI healthcare validation.

2. Aggregate diagnostic accuracy of 98.7% across all biomarker categories, validated against 1,000,000+ test cases with global population representation.

3. 87% correlation with confirmed clinical outcomes in longitudinal follow-up studies, demonstrating real-world effectiveness.

4. Continuous quality monitoring with quarterly model updates ensures adaptation to evolving clinical guidelines and emerging research.

5. Medical Advisory Board oversight provides clinical governance for all algorithmic decisions, ensuring appropriate integration of AI in healthcare.

The platform serves as a decision support tool designed to augment—not replace— professional medical judgment. Ongoing validation studies and transparent reporting remain central to our commitment to responsible AI deployment in healthcare.

References

[1] World Health Organization. (2023). WHO Model List of Essential In Vitro Diagnostics. WHO/MHP/HPS/EML/2023.02.

[2] National Institutes of Health. Clinical Laboratory Reference Values. National Library of Medicine. NBK532261.

[3] International Organization for Standardization. (2022). ISO 15189:2022 - Medical laboratories — Requirements for quality and competence.

[4] College of American Pathologists. (2024). Laboratory Accreditation Program Standards.

[5] European Federation of Clinical Chemistry and Laboratory Medicine. (2023). Guidelines for AI Applications in Laboratory Medicine.

[6] American Association for Clinical Chemistry. (2024). Reference Interval Database.

[7] Clinical and Laboratory Standards Institute. (2023). EP28-A3c: Defining, Establishing, and Verifying Reference Intervals.

Document Information
Version: 2.0 | Last Updated: November 2025 | Next Review: February 2026
Prepared by: PIYA AI Technical & Medical Team | Reviewed by: Medical Advisory Board

Disclosure: This validation documentation is published by Kantesti/PIYA AI. Medical Advisory Board members receive compensation for their advisory roles.

Contact: info@kantesti.net | +49 177 497 4039 | https://www.kantesti.net/medical-validation/

DOI: 10.5281/kantesti.2025.mv2