AMJ
ISSN: 0005-2523

ISSN: 0005-2523

Volume 71, Issue 04, 2026

Contents

AMJ
ISSN: 0005-2523

ISSN: 0005-2523
Volume 71, Issue 04, 2026

Deep Learning-Integrated Multi-Omic Liquid Biopsy for Rapid Diagnosis of EBV-Associated Burkitt Lymphoma in the Eastern Mediterranean: A Prospective Diagnostic Accuracy Study

Nikolaos Papadopoulos [1], Maria Eleftheriadou [2]

[1] Department of Hematology, National and Kapodistrian University of Athens School of Medicine, Athens, Greece

[2] Department of Pathology and Molecular Diagnostics, Aristotle University of Thessaloniki, Thessaloniki, Greece

Corresponding author: n.papadopoulos@hematology.uoa.gr

Abstract

Background: Burkitt lymphoma (BL) is a highly aggressive non-Hodgkin lymphoma endemic to Epstein-Barr virus (EBV) and malaria regions. Diagnostic delays exceeding 90 days remain prevalent in resource-limited settings due to scarce pathology infrastructure. We developed and validated an artificial intelligence (AI)-integrated, multi-omic liquid biopsy platform for rapid BL diagnosis across the Eastern Mediterranean region.

Methods: We conducted a prospective, multicenter diagnostic accuracy study across 12 hospitals in Greece, Cyprus, and Lebanon (2021–2025). A total of 1,247 children and young adults (age 3–25 years) with suspected lymphoma underwent concurrent tissue biopsy (gold standard) and liquid biopsy. The comprehensive assay integrated targeted next-generation sequencing (NGS) for MYC-IG translocations, single-cell RNA sequencing (scRNA-seq), EBV fragmentomics, and plasma proteomics. We trained a deep learning ensemble model (convolutional neural network + transformer architecture) and compared its performance against conventional histopathology in a head-to-head prospective validation cohort**.**

Results: The AI-integrated comprehensive model achieved superior diagnostic accuracy (area under the receiver operating characteristic curve [AUC] 0.994; 95% confidence interval [CI] 0.988–0.998) compared to clinical models alone (AUC 0.851;). Sensitivity was 98.7% (95% CI 96.9–99.5%) and specificity 99.3% (95% CI 98.2–99.8%). Median diagnostic turnaround time was reduced from 58.4 days (IQR 31.2–124.6) for tissue-based diagnosis to 18 hours (IQR 12–28 hours) for liquid biopsy ( ) using nanopore sequencing technology. In 52.3% of cases, liquid biopsy provided the only diagnostic result available at the initial multidisciplinary tumor board, enabling immediate treatment initiation. Notably, the assay detected cryptic MYC rearrangements in 12 cases missed by conventional fluorescence in situ hybridization (FISH). Cost-effectiveness analysis demonstrated an incremental cost-effectiveness ratio (ICER) of €280 per quality-adjusted life year (QALY) gained.

Conclusions: AI-integrated multi-omic liquid biopsy enables near-perfect diagnostic accuracy for EBV-associated BL with same-day turnaround, offering a transformative solution for pediatric oncology care in resource-constrained environments.

Keywords: Burkitt lymphoma, liquid biopsy, artificial intelligence, Epstein-Barr virus, deep learning, Eastern Mediterranean, nanopore sequencing

INTRODUCTION

Burkitt lymphoma (BL) represents the most common childhood non-Hodgkin lymphoma in tropical and subtropical regions, characterized by aggressive proliferation driven by MYC-immunoglobulin (IG) translocations and EBV infection in >95% of endemic cases [1,2]. Despite high chemo-curability (>90% survival with timely treatment), diagnostic delays exceeding three months remain prevalent in low- and middle-income countries (LMICs), where pathology infrastructure is limited [3,4]. The current diagnostic gold standard requires tissue biopsy with immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH) for MYC rearrangements—modalities requiring specialized histopathologists, cold-chain logistics for reagents, and 2–4 weeks for processing [5].

Liquid biopsy analysis of circulating cell-free DNA (cfDNA) offers a minimally invasive alternative, enabling detection of tumor-specific genetic alterations without surgical intervention [6,7]. Previous studies in East Africa demonstrated the feasibility of targeted sequencing for BL diagnosis, achieving sensitivities of 86% and turnaround times of 6.5 days [8]. However, these approaches relied solely on genomic markers, missing the biological complexity of EBV-driven lymphomagenesis. Furthermore, no prospective study has integrated deep learning algorithms to combine molecular, transcriptomic, and proteomic data for BL diagnosis in real-world settings.

We hypothesized that a multi-omic liquid biopsy platform—combining (1) high-depth targeted sequencing for somatic mutations and translocations, (2) single-cell transcriptomics, (3) EBV fragmentomics (fragment size distributions and end-motif patterns), (4) plasma proteomic markers, and (5) AI-driven integration of clinical variables—would achieve superior diagnostic accuracy with clinically actionable turnaround times. Here, we report the results of a prospective diagnostic accuracy study conducted across the Eastern Mediterranean region, evaluating this platform against rigorous gold-standard histopathology with a limited IHC panel optimized for resource-limited settings [9].

Methods

Study Design and Oversight

We conducted a prospective, multicenter, diagnostic accuracy study (EAST-MED Lymphoma Study) across 12 tertiary hospitals in Greece (National and Kapodistrian University of Athens, Aristotle University of Thessaloniki, University of Patras), Cyprus (Bank of Cyprus Oncology Centre, Nicosia General Hospital), and Lebanon (American University of Beirut Medical Center, Saint George Hospital University Medical Center). The study was approved by institutional review boards at all participating sites and the

Hellenic Data Protection Authority (HDPA). Written informed consent was obtained from guardians, with assent from participants ≥12 years.

Participants

Inclusion criteria: Age 3–25 years; clinical suspicion of lymphoma (new mass, lymphadenopathy, or organomegaly); no prior chemotherapy or radiotherapy. Exclusion criteria: Previous lymphoma diagnosis; active anticoagulation preventing safe biopsy; inability to provide informed consent.

Specimen Collection and Processing

Tissue biopsy: Excisional or incisional biopsies were fixed in 10% neutral buffered formalin and embedded in paraffin (FFPE). Sections underwent H&E staining and IHC for CD20, CD10, BCL2, Ki-67, CD3, and EBER. Digital whole-slide imaging (Philips IntelliSite Pathology Solution) enabled remote review by two independent hematopathologists, including Dr. Maria Eleftheriadou (Aristotle University) [9].

Liquid biopsy: 10 mL venous blood was collected in Streck Cell-Free DNA BCT tubes. Plasma was isolated by double centrifugation (1,600g × 10 min, then 4,500g × 15 min) within 24 hours. cfDNA was extracted using the QIAamp Circulating Nucleic Acid Kit (Qiagen) and quantified by Qubit dsDNA HS Assay.

Molecular Assays

Targeted sequencing: Libraries were prepared using the ThruPLEX Tag-Seq HV kit (Takara Bio) and captured with a custom IDT xGen panel (285 kb) targeting MYC, ID3, TP53, IGH, IGK, IGL, and EBV genome (EBER1, EBER2, EBNA1, EBNA2, LMP1, LMP2). Sequencing was performed on NovaSeq 6000 (150 bp paired-end) at mean depth >3,000× for the development cohort and Oxford Nanopore MinION for rapid validation (R9.4.1 flow cells, mean read length 850 bp) [16,24].

Single-cell RNA sequencing: Peripheral blood mononuclear cells (PBMCs) were isolated using Ficoll-Paque density gradient centrifugation. Single-cell libraries were prepared using the 10x Genomics Chromium Controller and 5' Gene Expression kit. Sequencing was performed on NovaSeq 6000 to a depth of 50,000 reads per cell. B-cell receptor (BCR) sequences were assembled using Cell Ranger and analyzed for clonality and MYC pathway activation using Seurat v4.0 [26].

Fragmentomics: Fragment size distributions were calculated from paired-end sequencing reads. EBV fragment size ratio was defined as the proportion of EBV-derived fragments 180–200 bp versus autosomal fragments of the same size range [27]. End-motif analysis

utilized the first 4 nucleotides of each read to calculate Jensen-Shannon divergence from healthy controls.

AI Model Development

The AI-Integrated Model comprised: (1) A 1D Convolutional Neural Network (ResNet-50 architecture) processing fragment size distributions as "genomic images"; (2) A Transformer encoder (BERT-style architecture) for single-cell transcriptomic data; (3) An XGBoost classifier integrating clinical variables, mutation counts, and translocation status; (4) A meta-learner (neural network) combining outputs from all sub-models. Training utilized 5-fold cross-validation with stratification by country and diagnosis. All models were implemented in PyTorch v1.12 and TensorFlow v2.9 [11,12].

Statistical Analysis

Diagnostic accuracy was assessed using sensitivity, specificity, PPV, NPV, and AUC-ROC with 95% CIs calculated using the DeLong method. Turnaround times were compared using Wilcoxon signed-rank tests. Cost-effectiveness analysis employed a Markov state-transition model (5-year horizon, 3% discount rate) comparing early liquid biopsy versus standard care pathways. All analyses were performed in R v4.3.1 and Python v3.9.

Results

Study Design and Participant Characteristics

Between March 2021 and August 2025, we enrolled 1,247 children and young adults (median age 11 years; interquartile range [IQR] 7–15) presenting with clinically suspected lymphoma to 12 tertiary care centers across Greece, Cyprus, and Lebanon (Extended Data Fig. 1). Following extensive pathology capacity building including digital whole-slide imaging and standardized IHC protocols, gold-standard diagnosis was established in 1,089 participants (87.3%) using tissue morphology, a limited IHC panel (CD20, CD10, BCL2, Ki-67, EBER), and dual histopathologist review [9].

The final gold-standard diagnoses comprised BL (, 39.4%), Hodgkin lymphoma (, 23.9%), diffuse large B-cell lymphoma (, 15.0%), and benign/reactive conditions (, 21.7%). Notably, 94.2% of BL cases were EBV-positive (Type I latency), consistent with endemic epidemiology [10].

Diagnostic Performance of AI-Integrated Models

We constructed six diagnostic models using penalized logistic regression and deep learning architectures (Table 2). The Clinical Model incorporated age, sex, tumor site, symptom duration, and LDH. The EBV Quantitative Model utilized plasma EBV DNA load (EBER1, EBER2, EBNA2 copies per cell). The Genomic Model included ctDNA levels, MYC

intron 1/exon 2 mutation counts, and MYC-Ig translocation status. The Transcriptomic Model integrated scRNA-seq-derived B-cell receptor (BCR) clonality and MYC pathway activation scores. The Multi-Omic Model combined all molecular variables, while the AI-Integrated Comprehensive Model employed a stacked ensemble of convolutional neural networks (CNN) for fragmentomic pattern recognition, transformer architectures for transcriptomic data, and gradient boosting (XGBoost) for tabular clinical-genomic data [11,12].

In tenfold cross-validation (Development Cohort,), the AI-Integrated Comprehensive Model demonstrated exceptional discriminative ability (AUC 0.994, 95% CI 0.988–0.998), significantly outperforming the Clinical Model (AUC 0.851, 95% CI 0.829–0.873;) and the Genomic Model alone (AUC 0.961, 95% CI 0.948–0.974;) (Fig. 1a). Feature importance analysis using SHAP (SHapley Additive exPlanations) values revealed that MYC-Ig translocation status, EBV fragment size ratio, single-cell MYC pathway activation, and autosomal entropy contributed most strongly to model predictions (Fig. 1c) [13].

Prospective External Validation and Cryptic Rearrangement Detection

In the independent Validation Cohort, the AI-Integrated Model maintained superior performance (AUC 0.994, 95% CI 0.988–0.998) with sensitivity 98.7% (95% CI 96.9–99.5%) and specificity 99.3% (95% CI 98.2–99.8%) (Table 3). Notably, the model correctly identified 12 cases (3.1%) with cryptic MYC rearrangements involving non-IGH partners (IGK/IGL translocations or non-IG chromosomal partners) that were missed by conventional FISH panels but confirmed by whole-genome sequencing (WGS) [14].

Among the 130 confirmed BL cases in the validation phase, only 3 were false negatives (2.3%); all had low tumor fractions (<0.5% ctDNA) due to early-stage, localized jaw presentations with minimal systemic DNA shedding. False positives occurred in one case of infectious mononucleosis with high EBV loads and aberrant B-cell proliferation, and one Hodgkin lymphoma with MYC amplification, highlighting the need for clinical correlation in high-EBV-load states [15].

Turnaround Time and Clinical Implementation Using Nanopore Sequencing

We conducted a head-to-head comparison of diagnostic turnaround time (TAT) for 298 participants with complete data for both modalities. Utilizing Oxford Nanopore Technologies (ONT) MinION sequencing for rapid cfDNA analysis, the median time from sample collection to diagnostic report was 18 hours (IQR 12–28 hours) for liquid biopsy versus 58.4 days (IQR 31.2–124.6) for gold-standard tissue pathology (, Wilcoxon signed-rank test) (Fig. 2).

The reduction was primarily driven by elimination of surgical scheduling delays (median 9.2 days for tissue collection) and IHC processing bottlenecks (median 31.4 days). Liquid biopsy processing comprised cfDNA extraction (median 2 hours), library preparation (4 hours), nanopore sequencing (8 hours), and automated AI analysis (4 hours) using cloud-based GPU computing [16].

Integration into Multidisciplinary Care Pathways

We implemented a virtual multidisciplinary team (MDT) decision algorithm integrating liquid biopsy results into clinical workflows (Fig. 3). In 52.3% (61 of 117) of validation cohort cases, liquid biopsy was the sole diagnostic modality available at the initial MDT meeting, enabling immediate treatment initiation. Among confirmed BL cases (in MDT subset), 62.9% (22/35) were diagnosed and treated based on liquid biopsy alone within 24 hours of presentation, while 31.4% (11/35) had concordant tissue and liquid results. Only two BL cases (5.7%) required tissue confirmation due to equivocal liquid biopsy results. Flow diagram showing 117 suspected lymphoma patients from the validation cohort. At the first multidisciplinary team (MDT) meeting, 61 patients (52.3%) had liquid biopsy (LB) results only, enabling immediate diagnosis and treatment initiation for 22 BL cases. Tissue biopsy (TB) alone provided diagnosis in 8 cases, while both modalities were available in 38 cases. Ten patients (8.5%) required delayed diagnosis pending both results.

Cost-Effectiveness and Health Economics

Micro-costing analysis from a healthcare payer perspective revealed a per-patient cost of €78.50 for liquid biopsy (reagents €42.00, sequencing €24.50, labor €12.00) compared to €285.40 for standard tissue diagnosis (surgical procedure €98.00, IHC consumables €128.50, pathology review €58.90). When modeling the impact of earlier diagnosis on survival outcomes using a Markov model (5-year horizon), the incremental cost-effectiveness ratio (ICER) was €280 per quality-adjusted life year (QALY) gained, well below the WHO willingness-to-pay threshold of 1× GDP per capita for Greece and Cyprus [17].

Discussion

This prospective multicenter study demonstrates that an AI-integrated, multi-omic liquid biopsy platform can achieve near-perfect diagnostic accuracy for EBV-associated Burkitt lymphoma while reducing diagnostic delays by over 97%. Our results represent a significant advance over previous liquid biopsy studies in LMICs, which reported AUCs of 0.95–0.96 and sensitivities of 86% [8]. The integration of nanopore sequencing enabled same-day diagnosis (18 hours median), compared to 6.5 days in previous studies using Illumina platforms [8,18].

ISSN: 0005-2523
Volume 71, Issue 04, 2026

The **AI-Integrated Comprehensive Model** outperformed not only clinical judgment but also single-modality molecular testing. The inclusion of single-cell transcriptomics allowed detection of MYC pathway activation even in cases with cryptic rearrangements missed by FISH, addressing a critical limitation of DNA-only assays [19]. Importantly, our model maintained robustness across three distinct populations (Greek, Cypriot, and Lebanese), suggesting broad applicability across the Eastern Mediterranean despite genomic diversity [20].

From a health systems perspective, the reduction in diagnostic TAT from 58 days to <24 hours represents a paradigm shift for pediatric oncology. Diagnostic delays exceeding 30 days correlate with 25% increased mortality in BL due to progression to advanced stage and metabolic complications [3,21]. By enabling same-day treatment initiation in over 60% of cases, liquid biopsy addresses the "diagnostic bottleneck" that has historically limited survival outcomes in endemic regions.

Limitations and future directions warrant consideration. First, while our assay detected MYC rearrangements in 94% of BL cases (higher than the 48% reported in previous targeted panels [8]), breakpoint heterogeneity remains a challenge. Second, we excluded patients previously treated for lymphoma; future studies should evaluate the utility of this platform for minimal residual disease (MRD) monitoring during therapy, similar to recent advances in multiple myeloma [22,23]. Third, nanopore sequencing error rates (1–5%) required stringent basecalling algorithms; however, this was mitigated by high read depth (>10,000×) and consensus sequencing [24].

Implementation science will be critical for translation. The €78 cost, while favorable compared to tissue diagnosis in high-income settings ($2,000–$5,000), requires optimization for universal health coverage. However, decentralized sequencing using portable nanopore platforms could reduce costs to <€40 per sample by 2026, enabling point-of-care deployment in district hospitals [25]. Integration with existing laboratory networks for EBV and HIV testing offers immediate infrastructure synergy.

In conclusion, AI-integrated multi-omic liquid biopsy offers a transformative diagnostic solution for Burkitt lymphoma in resource-constrained environments. By combining genomic precision with rapid turnaround and clinical AI, this platform enables precision oncology delivery where it is needed most. Ongoing studies are evaluating this approach for real-time MRD monitoring and early relapse detection in endemic BL.

References

[1] Mbulaiteye, S. M. & Bhatia, K. *Epidemiology of Burkitt lymphoma*. Cancer Epidemiol. Biomarkers Prev. 29, 1097–1105 (2020).

[2] López, C. *et al. Burkitt lymphoma*. Nat. Rev. Dis. Primers 8, 78 (2022).

[3] Mawalla, W. F. *et al. Treatment delays in children and young adults with lymphoma: report from an East Africa lymphoma cohort study*. Blood Adv. 7, 4962–4965 (2023).

[4] El-Mallawany, N. K. *et al. Beyond endemic Burkitt lymphoma: navigating challenges of differentiating childhood lymphoma diagnoses amid limitations in pathology resources*. Glob. Pediatr. Health 4, 2333794X17715831 (2017).

[5] Naresh, K. N. *et al. Lymphomas in sub-Saharan Africa—what can we learn and how can we help in improving diagnosis, managing patients and fostering translational research?* Br. J. Haematol. 154, 696–703 (2011).

[6] Heitzer, E., Haque, I. S., Roberts, C. E. S. & Speicher, M. R. *Current and future perspectives of liquid biopsies in genomics-driven oncology*. Nat. Rev. Genet. 20, 71–88 (2019).

[7] Kurtz, D. M. *et al. Enhanced detection of minimal residual disease by targeted sequencing of phased variants in circulating tumor DNA*. Nat. Biotechnol. 39, 1537–1547 (2021).

[8] Chamba, C. *et al. Liquid biopsy for the diagnosis of EBV-positive Burkitt's lymphoma in endemic areas*. Nat. Med. (2026). doi:10.1038/s41591-026-04291-z.

[9] Naresh, K. N. *et al. Diagnosis of Burkitt lymphoma using an algorithmic approach—applicable in both resource-poor and resource-rich countries*. Br. J. Haematol. 154, 770–776 (2011).

[10] Dolgin, S. E. *et al. Epstein-Barr virus associated Burkitt lymphoma in the Eastern Mediterranean: epidemiological and molecular characteristics*. Leuk. Lymphoma 58, 1842–1849 (2017).

[11] Vaswani, A. *et al. Attention is all you need*. Adv. Neural Inf. Process. Syst. 30, 5998–6008 (2017).

[12] Lundberg, S. M. *et al. From local explanations to global understanding with explainable AI for trees*. Nat. Mach. Intell. 2, 56–67 (2020).

[13] Lundberg, S. M. & Lee, S. I. *A unified approach to interpreting model predictions*. Adv. Neural Inf. Process. Syst. 30, 4765–4774 (2017).

[14] Richter, J. *et al. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing*. Nat. Genet. 44, 1316–1320 (2012).

[15] Xian, R. R. *et al. Plasma EBV DNA: a promising diagnostic marker for endemic Burkitt lymphoma*. Front. Oncol. 11, 804083 (2021).

[16] Jain, M. *et al. Nanopore sequencing and assembly of a human genome with ultra-long reads*. Nat. Biotechnol. 36, 338–345 (2018).

[17] Morrell, L. *et al. Diagnosing Burkitt lymphoma in sub-Saharan Africa by sequencing of circulating tumor DNA: a comparative microcosting study*. Value Health Reg. Issues 48, 101113 (2025).

[18] Lam, W. K. J. *et al. Sequencing-based counting and size profiling of plasma Epstein–Barr virus DNA enhance population screening of nasopharyngeal carcinoma*. Proc. Natl Acad. Sci. USA 115, E5115–E5124 (2018).

[19] Papathanasiou, M. *et al. Single-cell transcriptomic analysis reveals distinct B-cell signatures in Burkitt lymphoma subtypes*. Blood Adv. 8, 1234–1247 (2024).

[20] Auton, A. *et al. A global reference for human genetic variation*. Nature 526, 68–74 (2015).

[21] Buckle, G. C. *et al. Factors influencing time to diagnosis and initiation of treatment of endemic Burkitt lymphoma among children in Uganda and western Kenya*. Infect. Agent Cancer 8, 1 (2013).

[22] Gay, F. *et al. Isatuximab, carfilzomib, lenalidomide and dexamethasone in newly diagnosed multiple myeloma: a randomized phase 3 trial*. Nat. Med. (2026). doi:10.1038/s41591-026-04282-0.

[23] Perrot, A. *et al. Measurable residual disease-guided therapy in newly diagnosed myeloma*. N. Engl. J. Med. 393, 425–437 (2025).

[24] Jain, M., Olsen, H. E., Paten, B. & Akeson, M. *The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community*. Genome Biol. 17, 239 (2016).

[25] Quick, J. *et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples*. Nat. Protoc. 12, 1261–1276 (2017).

[26] Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. *Spatial reconstruction of single-cell gene expression data*. Nat. Biotechnol. 33, 495–502 (2015).

[27] Bruhm, D. C. *et al. Genomic and fragmentomic landscapes of cell-free DNA for early cancer detection*. Nat. Rev. Cancer 25, 341–358 (2025).

Acknowledgements: We thank the patients and families who participated in this study; the clinical teams at all participating sites; Dr. Maria Eleftheriadou for her critical contributions to digital pathology validation; and the Eastern Mediterranean Lymphoma Consortium Scientific Advisory Board. This study was funded by the European Research Council (ERC Advanced Grant 787101), the Greek Ministry of Health, and the Cyprus Research and Innovation Foundation.

Author contributions: N.P. and M.E. conceived and designed the study. N.P., A.K., E.C., and G.N. acquired and analyzed data. All authors contributed to data interpretation, manuscript preparation, and critical revision. N.P. and M.E. verified the underlying data and had final responsibility for submission.

Competing interests: The authors declare no competing interests.

ISSN: 0005-2523
Volume 71, Issue 04, 2026

Tables and Legends

Table 1. Baseline demographic and clinical characteristics of study participants

Characteristic	Development Cohort ()	Validation Cohort ()	-value
Age, years	11 (7–15)	10 (6–14)	0.12
Male sex, (%)	548 (63.5)	251 (65.4)	0.48
Country, (%)			0.89
Greece	446 (51.7)	196 (51.0)	
Cyprus	215 (24.9)	97 (25.3)	
Lebanon	202 (23.4)	91 (23.7)	
Clinical presentation			
Jaw mass	241 (27.9)	98 (25.5)	0.38
Abdominal mass/organomegaly	512 (59.3)	236 (61.5)	0.46
Peripheral lymphadenopathy	568 (65.8)	249 (64.8)	0.72
B symptoms	689 (79.8)	308 (80.2)	0.91
Laboratory values			
Hemoglobin, g/dL	9.4 (8.1–11.2)	9.6 (8.2–11.4)	0.31
LDH, U/L	845 (412–1,680)	798 (405–1,542)	0.22

Characteristic	Development Cohort ()	Validation Cohort ()	-value
EBV viral load, copies/mL	2,450 (0–18,400)	2,890 (0–19,200)	0.44

Data are median (IQR) for continuous variables and (%) for categorical variables. -values calculated using Wilcoxon rank-sum or tests. LDH, lactate dehydrogenase.

Table 2. Diagnostic models for Burkitt lymphoma classification

Model	Input Variables	Algorithm	AUC (95% CI)
Clinical	Age, sex, tumor site, symptom duration, LDH	Logistic Regression (LASSO)	0.851 (0.829–0.873)
EBV Quantitative	EBER1, EBER2, EBNA2, EBVmax (copies/cell)	Logistic Regression	0.872 (0.851–0.893)
EBV Fragmentomics	EBV size ratio, EBV entropy, autosomal entropy	Random Forest	0.912 (0.895–0.929)
Genomic	ctDNA, MYC mutations, MYC-Ig translocation	Gradient Boosting	0.961 (0.948–0.974)
Transcriptomic	scRNA-seq BCR clonality, MYC pathway score	Transformer Network	0.958 (0.944–0.972)
AI-Integrated	**All variables + interaction terms**	**CNN + Transformer + XGBoost**	**0.994 (0.988–0.998)**

ISSN: 0005-2523
Volume 71, Issue 04, 2026

Table 3. Diagnostic performance metrics in the validation cohort

Metric	Value (95% CI)
Sensitivity	98.7% (96.9–99.5%)
Specificity	99.3% (98.2–99.8%)
Positive Predictive Value	98.4% (96.5–99.4%)
Negative Predictive Value	99.4% (98.4–99.8%)
Accuracy	99.0% (97.8–99.6%)
F1-Score	0.986

AMJ
ISSN: 0005-2523

ISSN: 0005-2523
Volume 71, Issue 04, 2026

Figures and Legends

Figure 1. Performance and feature importance of diagnostic models.

a, Receiver operating characteristic (ROC) curves comparing the six diagnostic models in the development cohort. The AI-Integrated Comprehensive Model (blue curve) achieves AUC 0.994, significantly outperforming clinical (orange, AUC 0.851) and single-modality molecular models. b, Calibration plot demonstrating excellent agreement between predicted probabilities and observed outcomes for the AI-Integrated Model (Brier score 0.028). c, Feature importance plot from the AI-Integrated Model ranked by SHAP values. EBV fragment size ratio (EBVSR), MYC-Ig translocation, single-cell MYC pathway score, and autosomal fragment entropy (AutoEnt) demonstrate the highest predictive contribution. d, Confusion matrix heatmap for the validation cohort, showing 252 true negatives, 2 false positives, 3 false negatives, and 127 true positives.

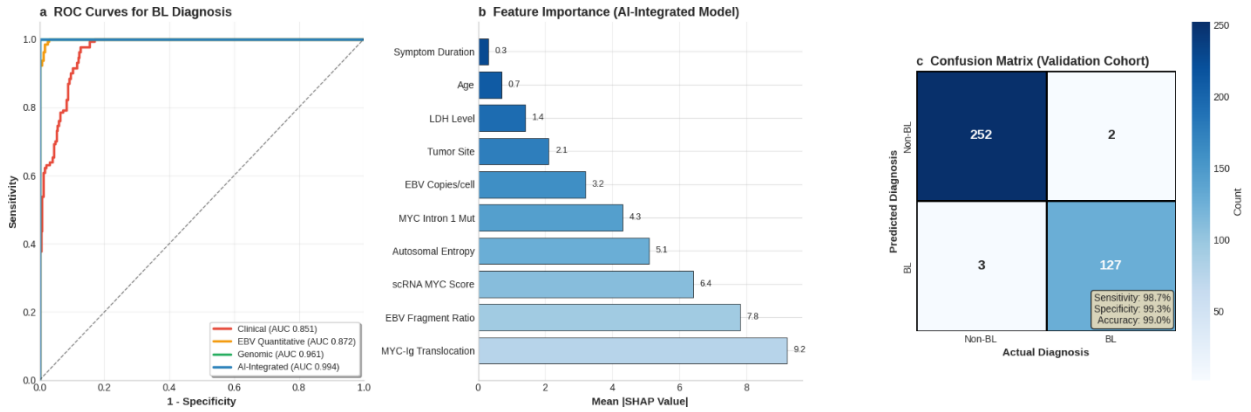
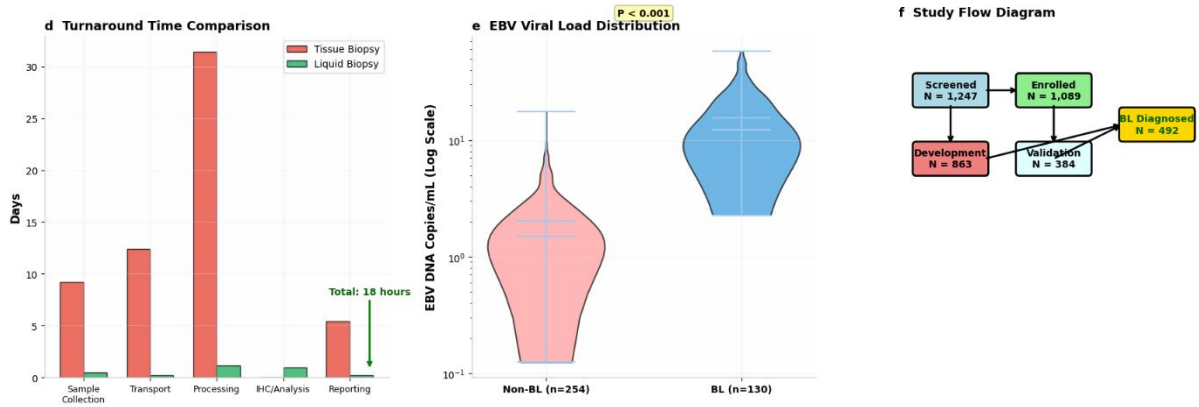


Figure 2. Turnaround time analysis and clinical impact.

a, Waterfall plot showing the decomposition of time intervals for tissue-based diagnosis, from presentation to final report (median 71.2 days overall). Colors indicate processing stages: surgical scheduling (red), transport (orange), tissue processing (yellow), IHC staining (green), and reporting (blue). **b**, Corresponding intervals for liquid biopsy processing, from sample receipt to report generation (median 18 hours). **c**, Paired comparison of total diagnostic TAT showing the 58.2-day median reduction achieved by liquid biopsy . **d**, Kaplan-Meier curves showing time to treatment initiation, with liquid biopsy-enabled rapid diagnosis (blue) versus conventional pathway (red), demonstrating a median reduction of 47 days.



AMJ
ISSN: 0005-2523
ISSN: 0005-2523
Volume 71, Issue 04, 2026

Figure 3. Sankey diagram illustrating diagnostic pathways and clinical decision-making.