

Title: A CSF ctDNA-based molecular panel for minimally invasive diagnosis of primary CNS lymphoma

Authors: Haoyi Li^{1,2†}, Huizhi Wang^{1,2†}, Shaoping Shen^{1,3†}, Bing Li^{4†}, Jinhao Zhang^{2†}, Hao Duan^{5†}, Ming Li¹, Xinze Lv⁴, Gehong Dong⁶, Leilei Yuan⁷, Zairan Wang¹, Mingxiao Li⁸, Linxiao Li¹, Tengfei Zhu⁴, Qian Liu², Lin Sang⁹, Dabiao Zhou^{1,10}, Ling Qin¹¹, Song Lin^{1,2,10}, Yonggao Mou⁵, Zheng Zhao^{2,12*}, Fangang Meng^{1,2,13*}, Feng Li^{14*}, Xiaohui Ren^{1,2,10,15*}

Affiliations:

¹Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing, China.

²Beijing Neurosurgical Institute, Capital Medical University, Beijing, China.

³Department of Neurosurgery, The First Affiliated Hospital of the Naval Medical University, Shanghai, China.

⁴Burning Rock Biotech, Guangdong, China.

⁵Department of Neurosurgery, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou, China.

⁶Department of Pathology, Beijing Tiantan Hospital, Capital Medical University, Beijing, China.

⁷Department of Nuclear Medicine, Beijing Tiantan Hospital, Capital Medical University, Beijing, China.

⁸Department of Neurosurgery, China-Japan Friendship Hospital, Beijing, China.

⁹Department of Neurosurgery, Beijing Fengtai Hospital, Beijing, China.

¹⁰China National Clinical Research Center for Neurological Diseases, Beijing, China.

¹¹Department of Infectious Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China.

¹²Chinese Glioma Genome Atlas Network (CGGA) and Asian Glioma Genome Atlas Network (AGGA), Beijing, China.

¹³Chinese Institute for Brain Research, Beijing, China.

¹⁴Department of Neurosurgery, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China.

¹⁵Lead contact

†Contributed equally

*Corresponding author: xiaohui ren@mail.ccmu.edu.cn (X. Ren), lidoc@163.com (F. Li), fgmeng@ccmu.edu.cn (F. Meng), zhaozheng0503@ccmu.edu.cn (Z. Zhao)

Abstract

Background: Primary central nervous system lymphoma (PCNSL) progresses rapidly and requires prompt diagnosis, yet tissue biopsy is invasive and may be infeasible in high-risk patients. Reliable minimally invasive diagnostic strategies are unmet.

Methods: In this multicenter diagnostic study, 174 patients with radiologically suspected PCNSL, including 103 with pathologically confirmed PCNSL, were prospectively enrolled across two Chinese centers. Available tumor tissue (TIS), cerebrospinal fluid (CSF) and plasma specimens were sequenced. For comparison, TIS sequencing data from 54 patients with grade 4 gliomas were retrieved from the CGGA database. Publicly available TIS and CSF sequencing data from 29 patients with PCNSL and 28 with grade 4 gliomas at a US center were included for validation. A coarse-to-fine machine-learning framework was applied to develop TIS-based molecular panels. Mutational concordance between matched TIS and liquid biopsy samples was assessed, and panel performance was further validated using CSF circulating tumor DNA (ctDNA).

Results: We developed and validated three TIS-based diagnostic panels capable of distinguishing PCNSL from grade 4 glioma. Mutational concordance analysis indicated CSF as the most reliable liquid biopsy source for molecular diagnosis. The random forest-optimized eight-gene panel (N8-RF-PCNSL) enabled CSF ctDNA-based diagnosis of PCNSL (sensitivity $\geq 79\%$, specificity: 100%, AUC ≥ 0.90) and identified a subset of *MYD88*^{wt} PCNSL with 100% specificity across the two Chinese validation cohorts. However, diagnostic performance was limited in the US validation set. Combining N8-RF-PCNSL with CSF interleukin-10 testing further increased sensitivity to 86%-93% while maintaining 100% specificity. This synergistic approach also provided supportive molecular evidence for PCNSL in over 60% of biopsy-infeasible patients, who subsequently demonstrated favorable radiographic responses.

Conclusion: We developed and validated a CSF ctDNA-based eight-gene molecular panel (N8-RF-PCNSL) providing minimally invasive molecular support for PCNSL diagnosis. This work supports a practical CSF-first diagnostic strategy to facilitate timely and accurate decision-making in neuro-oncology.

Keywords: primary CNS lymphoma, circulating tumor DNA, diagnosis, machine learning, interleukin-10 (IL-10)

Introduction

Primary CNS lymphoma (PCNSL) is a rare but aggressive extra-nodal non-Hodgkin lymphoma characterized by heterogeneous clinical outcomes in immunocompetent patients [1, 2]. Differentiating PCNSL from radiographically similar entities, such as high-grade gliomas, cerebral metastases, or inflammatory demyelination, remains a significant challenge in clinical practice [3]. Given its rapid clinical progression, timely and accurate diagnosis is essential to secure the optimal treatment window and improve survival outcomes.

Tumor tissue (TIS) biopsy has been the gold standard for PCNSL diagnosis [2, 4]. However, its routine clinical application is limited by procedural invasiveness, high cost and interobserver variability. As an invasive procedure, TIS biopsy has inherent risks, such as life-threatening hemorrhage in patients with deep-seated lesions, and may be poorly tolerated in elderly or medically fragile individuals. Furthermore, diagnostic accuracy may be compromised by pre-biopsy corticosteroid exposure [5-9]. These limitations have stimulated increasing interest in minimally invasive liquid biopsy strategies, particularly cerebrospinal fluid (CSF)-based biomarkers, for PCNSL diagnosis.

Conventional CSF-based diagnostic methods, including cytology and flow cytometry, are routinely employed but demonstrate limited sensitivity in identifying PCNSL [10]. Recent studies, along with the guidelines from the European Hematology Association-European Society for Medical Oncology (EHA-ESMO), have recommended testing for CSF interleukin-10 (IL-10), either alone or in conjunction with *MYD88* L265P mutation assessment, to enhance diagnostic discrimination [11-14]. Nevertheless, its clinical adoption is still constrained by substantial heterogeneity in IL-10 thresholds and the lack of standardization across laboratories [15].

Emerging evidence shows that CSF circulating tumor DNA (ctDNA) outperforms plasma (PLA) in genomic fidelity and permits sensitive detection of PCNSL hallmark mutations [16-22]. Among these, the canonical *MYD88* L265P mutation is the most extensively investigated and is recognized as a highly specific genetic hallmark for PCNSL [16-22]. However, *MYD88* L265P-based assays show widely variable sensitivity (25.0%-90.0%) [20, 23], likely due to the constraints of single-mutation

detection and the biological heterogeneity of PCNSL. Notably, approximately 30%-40% of PCNSLs exhibit molecularly atypical profiles (*MYD88*^{wt}) [24, 25], creating a critical diagnostic blind spot when relying exclusively on this mutation testing. To address this gap, a multiparametric CNS lymphoma sequencing panel was recently proposed and demonstrated moderate-to-high diagnostic performance, but its applicability for reliably identifying *MYD88*^{wt} PCNSL remains inconclusive [19]. Additionally, this panel incorporates numerous genetic predictors (26 biomarkers), potentially increasing financial and technical burdens, thereby limiting its clinical applicability in low-resource settings. Therefore, there is an urgent need for a streamlined, cost-effective, and clinically deployable targeted biomarker panel tailored for CSF ctDNA-based diagnostics of PCNSL [19, 24, 25].

In this study, we sought to develop a streamlined CSF ctDNA-based molecular panel capable of biopsy-free diagnosis of PCNSL. To address the most common diagnostic dilemma in clinical practice, the differentiation of PCNSL from grade 4 glioma, we first trained multiple machine-learning (ML) algorithms to derive TIS-based minimal gene panels with optimal discriminative performance. We then assessed the feasibility of liquid biopsy by analyzing mutational concordance across paired TIS, CSF, and PLA samples. Subsequently, we evaluated the diagnostic performance and broad clinical utility of the optimized panel for CSF ctDNA-based PCNSL classification (including *MYD88*^{wt} PCNSL). Finally, we investigated the potential synergistic diagnostic value of combining the gene panel with CSF IL-10 for improved identification of PCNSL.

Methods

Patient enrollment and genomic sequencing

This study included data from cohorts in China and the United States. The Chinese cohorts were sourced from Beijing Tiantan Hospital (BJTTH) and Beijing Fengtai Hospital (BJFTH), both of which were conducted prospectively. Ethical approvals were acquired from the respective institutional review boards (approval numbers: KY2024-093-02 for BJTTH and P20210005 for BJFTH), and written informed consent was secured from all participants. The US cohort comprised publicly accessible data from Memorial Sloan Kettering Cancer Center (MSKCC), retrieved via the cBioPortal database

[26, 27]. The MSKCC cohort, originally launched as a prospective study, served as an external validation dataset for post hoc analysis. This study adhered to the principles outlined in the Declaration of Helsinki.

In the BJTTH cohort (**Figure 1A**), 133 immunocompetent patients with radiologically suspected PCNSL were prospectively and consecutively enrolled in the Department of Neurosurgery between June 2024 and December 2025. At initial admission, patients underwent CSF sampling by lumbar puncture and peripheral venous blood collection before biopsy. Approximately 10 mL of each liquid sample was collected per patient. In total, 130 CSF samples and 76 plasma (PLA) samples were available, including 61 paired CSF-PLA samples. Biopsies were subsequently performed. Of the 133 patients, 91 were eligible for and underwent navigation-assisted stereotactic biopsy for histopathological confirmation. Ultimately, 76 were pathologically confirmed to have PCNSL while 15 were diagnosed with non-PCNSL entities. The non-PCNSL group consisted of glioblastoma (N = 4), IDH-mutant grade 4 astrocytoma (N = 1), anaplastic astrocytoma (N = 2), focal cerebral inflammatory lesions (N = 2), demyelinating disease (N = 2), brain metastasis (N = 1), gliosis (N = 1), cerebral infarction (N = 1), and craniopharyngioma (N = 1). The remaining 42 of the 133 patients with radiologically suspected PCNSL at BJTTH did not undergo surgical biopsy during the initial diagnostic evaluation because of surgical contraindications, high anesthetic risk, or refusal of invasive diagnostic procedures (**Table S1**). All 42 patients underwent evaluation by the institutional multidisciplinary team (MDT). Each patient was assessed individually based on clinical characteristics (e.g., presenting neurological manifestations, performance status), radiological features (e.g., contrast-enhancement pattern, diffusion restriction), systemic staging (e.g., whole-body ^{18}F -FDG PET/CT or PET/MRI), and CSF findings (e.g., CSF ctDNA, CSF IL-10 level, cytology). For patients with CSF molecular findings supportive of PCNSL, the MDT rendered a molecular diagnosis of PCNSL following comprehensive evaluation. Given the patients' clinical vulnerability, the diagnostic uncertainty resulting from the absence of histopathological confirmation, the urgent need for disease control, and the potential risks of conventional chemotherapy, prompt intervention with a targeted therapy-based regimen was considered appropriate to balance the potential clinical benefits against treatment-related risks in the setting of rapidly progressive disease and high tumor burden. The ethics approval for the diagnostic

study covered biospecimen collection and molecular testing only; this treatment regimen was not assigned by the study protocol but was administered as individualized clinical care following multidisciplinary evaluation and with the patients' informed consent. Further invasive diagnostic procedures were pursued whenever clinically necessary and tolerable, particularly in patients without CSF findings supportive of PCNSL.

Available TIS, CSF, and PLA specimens were sequenced using a capture-based targeted panel that comprises 112 genes associated with systemic diffuse large B-cell lymphoma (DLBCL, **Table S2**) [28]. The methods for DNA isolation and DNA sequencing were as previously described [29]. In brief, TIS DNA was extracted from formalin-fixed, paraffin-embedded (FFPE) tumor tissues using QIAamp DNA FFPE tissue kit (Qiagen, Hilden, Germany). CtDNA was extracted from CSF and PLA using QIAamp Circulating Nucleic Acid Kit (Qiagen, Hilden, Germany). DNA fragments between 200-400 bp were then purified using Agencourt AMPure XP Kit (Beckman Coulter, CA, USA). The panel spanning 314 kilobases of the human genome was used for target capture, with an average sequencing depth of 1000× for TIS/CSF and 10000× for PLA samples. Indexed samples were sequenced on NextSeq 500 (Illumina, Inc., CA, USA) with paired-end reads.

All procedures were performed at a commercial clinical laboratory (Burning Rock Biotech) accredited by the College of American Pathologists and certified by the Clinical Laboratory Improvement Amendments. The sequencing data were first mapped to the reference human genome (hg19) using Burrows-Wheeler Aligner version 0.7.10. Local alignment optimization, duplication marking and variant calling were performed using Genome Analysis Toolkit version 3.2, and VarScan version 2.4.3, respectively. Sequencing data of corresponding white blood cells were used to filter out germline variants and clonal haematopoiesis. For TIS and CSF samples, single nucleotide variations (SNVs) were retained when the mutant allele frequency (MAF) was $\geq 1.0\%$ and at least eight supporting reads were present; insertion/deletion variations (Indels) were retained when the MAF was $\geq 0.5\%$ and at least five supporting reads were present. For PLA samples, SNVs were retained when the MAF was $\geq 0.1\%$ and supported by at least eight reads, whereas Indels were retained when the MAF was $\geq 0.05\%$ and supported by at least two reads. Liquid samples were considered ctDNA-sufficient only if they met all of the following criteria: CSF/whole blood sample volume ≥ 8 mL; total nucleic acid yield $\geq$

15 ng for CSF and ≥ 30 ng for PLA; pre-library DNA ≥ 300 ng; coverage uniformity $\geq 90\%$; $\geq 80\%$ of bases with Q30 quality; and $\geq 90\%$ homozygous concordance with matched normal. Only samples that passed all quality control thresholds were included in mutation calling and downstream analyses, thereby ensuring robust and clinically meaningful results.

For comparison, clinicopathological data from 54 patients with treatment-naïve primary grade 4 glioma were retrieved from the clinical record system of BJTTH, and corresponding TIS-based whole-exome sequencing (WES) data were extracted from the Chinese Glioma Genome Atlas (CGGA) database [30].

In the BJFTH cohort (**Figure 1A**), we enrolled 41 immunocompetent patients who were radiologically suspected of having PCNSL in the Department of Neurosurgery between January and June 2025. CSF sample collection and targeted panel sequencing were performed using the same procedures as those employed in the BJTTH cohort [28]. All patients underwent navigation-assisted stereotactic biopsy, and histopathological examination confirmed 27 and 14 cases of PCNSL and non-PCNSL, respectively. The non-PCNSL group consisted of glioblastoma (N = 10), IDH-mutant grade 4 astrocytoma (N = 1), demyelinating disease (N = 2), and mesenchymal-origin malignant brain tumor (N = 1).

In the MSKCC cohort (**Figure 1A**), 57 eligible cases were enrolled for analysis after screening (**Table S3–4**). This cohort included 29 PCNSL cases (25 TIS and 23 CSF samples) [26] and 28 grade 4 glioma cases (13 TIS and 28 CSF samples) [27]. Clinicopathological information and targeted sequencing data from TIS and CSF samples were obtained from online public resources. Sequencing had been performed using either the MSK-HemePACT targeted panel [26] or the MSK-IMPACT assay, as previously described [31, 32].

Given the heterogeneity of the sequencing platforms used across cohorts, sequencing data were harmonized before feature selection, model development and external validation. The characteristics of the sequencing platform used in each cohort are summarized in **Table S5**. Cross-platform harmonization was performed at the gene-feature level, with the 112-gene panel (Burning Rock Biotech) as the reference gene set. The WES data from CGGA were restricted to genes included in

the reference gene set. For the MSKCC cohorts, only genes available in both the reference gene set and the sequencing data generated by MSK-HemePACT or MSK-IMPACT were evaluated. Eligible alterations were restricted to protein-altering coding variants. Each evaluable gene was encoded as mutated when at least one eligible somatic variant was reported and as non-mutated when no eligible variant was reported in the harmonized mutation dataset. Multiple eligible alterations within the same gene were represented by a single positive gene-level feature. Specific mutation types, the number of alterations within a gene and variant allele fractions were not used as model features. This procedure ensured that feature selection, model development, and validation were based on consistently defined gene-level mutation features.

To facilitate interpretation of the subsequent analyses, we clarified the composition and analytical role of each cohort (**Figure 1B**). TIS sequencing data from the BJTTH (64 PCNSL and 54 grade 4 glioma samples) were designated as the TIS-based training set for panel development, while publicly available tumor sequencing data from the MSKCC (25 PCNSL and 13 grade 4 glioma samples) served as a TIS-based external validation set. CSF ctDNA sequencing data from BJTTH, BJFTH and MSKCC were used to evaluate the diagnostic accuracy and clinical applicability of the proposed panels. Specifically, the diagnostic performance in distinguishing patients with PCNSL, including *MYD88*^{wt} PCNSL, from those without was assessed using CSF ctDNA data from BJTTH and BJFTH (CSF-based BJTTH/BJFTH validation cohort). The ability to differentiate patients with PCNSL from those with grade 4 glioma was evaluated using CSF ctDNA data from the MSKCC (CSF-based MSKCC validation set). Finally, an additional 42 radiologically suspected PCNSL patients from BJTTH constituted a Biopsy-infeasible real-world cohort to preliminarily assess the clinical applicability of the combined CSF-based molecular and IL-10 diagnostic strategy.

Genetic predictor selection and diagnostic panel development and validation

A coarse-to-fine feature-selection and model-development strategy was applied to the entire TIS-based training set without internal cross-validation. First, univariate analysis was conducted using Fisher's exact test to compare the mutation frequency of each gene detected in the TIS-based training set in patients with PCNSL vs. grade 4 glioma. Candidate genetic biomarkers (i.e., differentially mutated genes, DMGs) with a mutation frequency of $\geq 5\%$ and a *P* value of ≤ 0.05 were

selected for further analysis. Next, the preselected genetic predictors (i.e., DMGs) were optimized using the machine learning (ML) algorithms, Random Forest (RF), Support Vector Machine (SVM), and XGBoost (XGB), and then ranked in descending order based on their importance scores, while the cumulative area under the receiver operating characteristic curve (AUC) was continuously calculated with the predictors added consecutively. All models were trained using R packages with fixed parameters: i) RF: $n_{tree} = 500$, $m_{try} = \sqrt{n_{features}}$ (randomForest); ii) SVM: radial kernel, $C = 1$ (e1071); iii) XGB: $\eta = 0.3$, $max_depth = 6$, $n_{rounds} = 100$ (xgboost). Finally, the panels with optimized combination of genetic biomarkers were determined by identifying the cumulative AUC's inflection point and achieving the AUC's best performance.

External validation was conducted using the TIS-based external validation set. Confusion metrics of sensitivity, specificity, false positive rate and false negative rate were reported. To assess the reliability of three ML-generated models, we calculated: i) Log-loss, which penalizes inaccurate or overconfident predictions; ii) Brier score, which measures the squared deviation between predicted probabilities and observed labels. These metrics were computed on both the TIS-based training and external validation sets using the `LogLoss()` function from `MLmetrics` (v1.1.1) and the `BrierScore()` function from `DescTools` (v0.99.47).

CSF IL-10 levels

Given the lack of consensus on a standardized CSF IL-10 threshold for discriminating PCNSL across institutions, we used 5 pg/mL as the predefined cutoff in this study. This value corresponds to the lower reportable limit of the central laboratory assays used at both BJTTH and BJFTH, and was therefore adopted as an institutionally harmonized laboratory threshold rather than an empirically optimized cutoff derived from the present datasets.

Web tool development

To facilitate research use, an interactive web application was developed using the Shiny R package (version 1.14.0). The final selected model was locked before web implementation, and the application used the same feature definitions, feature-encoding rules, fitted ML model and classification threshold as those used in the validation analyses.

Statistical analysis

All analyses were performed using R software (version 4.2.3). Clinical characteristics were summarized descriptively, with continuous variables presented as medians with ranges and categorical variables presented as counts and percentages. DMGs were preselected during exploratory screening using Fisher's exact tests, and multiple-testing correction with Benjamini-Hochberg procedure was further conducted across all tested genes to obtain false-discovery rate (FDR)-adjusted *P* values. Unsupervised clustering of gene-level mutational profiles was conducted using principal component analysis (PCA) and K-means clustering. Mutational concordance between sample types was evaluated using Cohen's Kappa coefficient and sensitivity analysis. Sensitivity was calculated as the proportion of mutated genes detected in TIS samples that were also identified in the corresponding CSF or PLA samples. Inter-group differences were assessed using Wilcoxon rank-sum test. Diagnostic performance was evaluated using the AUC, sensitivity, and specificity, with corresponding 95% confidence intervals (CIs). AUCs were compared using DeLong's test. A two-sided *P* value of less than 0.05 was considered statistically significant.

Results

Patient characteristics

The clinical characteristics of patients with PCNSL from the BJTTH cohort are presented in **Figure 2A**. Among the 64 TIS samples (64/76, 84.2%) available for targeted sequencing, 60 (60/64, 93.8%) exhibited detectable mutations. Regarding the 73 CSF samples (73/76, 96.1%), mutations were identified in 63 cases (63/73, 86.3%). Although all 76 PLA samples (76/76, 100%) were available for targeted sequencing, mutations were detected in only 40 cases (40/76, 52.6%). Collectively, these findings suggest that (i) TIS remains the gold standard for mutation detection, considering its highest detection rate among all sample types, and (ii) CSF is superior to PLA for ctDNA detection.

Clinical features were generally comparable across cohorts and subgroups. For instance, the median age of the 76 patients (56.6% males) with PCNSL in the BJTTH cohort was 63 years (range: 27-79), compared to 68 years (range: 39-81) in the BJFTH cohort (59.3% males) and 69 years (range: 32-90)

in the MSKCC cohort (62.1% males). Detailed baseline data for patients with PCNSL, non-PCNSL, and grade 4 glioma across the three cohorts are summarized in **Table 1**.

CSF IL-10 positivity (> 5 pg/mL), however, varied substantially between cohorts (BJTTH: 40.8%; BJFTH: 81.5%) but remained absent in non-PCNSL controls in both cohorts (0/15 and 0/14, respectively; **Table 1**). Cytologic positivity rates were consistently low (BJTTH: 19.7%; BJFTH: 18.5%; **Table 1**). These findings underscore the need for a standardized, genetics-informed CSF-based diagnostic strategy to enhance diagnostic yield while preserving specificity and supporting clinical decision-making.

Insufficiency of the capture-based targeted panel in differential diagnosis

The capture-based targeted panel comprising 112 hotspot-mutated probes was originally designed for molecular profiling of systemic DLBCL. We therefore initially assessed its utility for PCNSL identification. PCA revealed that this panel could not fully distinguish patients with PCNSL from those with grade 4 glioma, as a subset of cases exhibited overlap (**Figure S1A**). K-means clustering analysis quantitatively demonstrated that 16 out of 118 (13.6%) cases were misclassified when distinguishing between patients with PCNSL and those with grade 4 glioma (**Figure S1B**). These results underscore the need for a more specific molecular panel for PCNSL diagnosis.

Development and validation of optimized molecular panels

Before modeling, we briefly assessed the comparability of sequencing data across analytical datasets after cross-platform gene-level harmonization. We observed that the distribution patterns of gene-level mutation detection rates in PCNSL and grade 4 glioma in the TIS-based training set were similar to those in the TIS-based external validation set, suggesting reasonable comparability between these datasets (**Table S6** and **Figure S2**). Similar comparability was also observed between the CSF-based BJTTH and BJFTH validation cohorts, but not between the CSF-based BJTTH/BJFTH validation cohorts and the CSF-based MSKCC validation set (**Table S7** and **Figure S2**).

As the first step of our coarse-to-fine feature selection strategy, we performed univariate gene-level exploratory screening to identify candidate PCNSL-enriched genetic biomarkers in TIS-based training set. Nineteen DMGs met the exploratory preselection criteria of a mutation frequency of $\geq 5\%$ and a

nominal P value of ≤ 0.05 . Of these, 12 remained significant after Benjamini-Hochberg correction across all tested genes (FDR-adjusted $P < 0.05$; **Table S8**). These 19 candidate DMGs were used for subsequent model development. In addition, the recurrently mutated genes *MYD88* (73% vs. 0%) and *PIM1* (69% vs. 0%) were highlighted (**Figure 2B** and **Figure S3**).

Utilizing the 19 DMGs, we employed three ML algorithms, including SVM, RF, and XGB, to optimize genetic biomarker combinations that demonstrate specificity for PCNSL diagnosis. Each ML platform ranked the genetic biomarkers in descending order of predictability. A sequential forward selection strategy, illustrated using line charts, was employed to balance the trade-off between model performance (AUC on the right axis) and the number of genetic biomarkers included. In RF and SVM algorithms, the line chart displayed a sharp increase followed by a gradual ascent (**Figure 2C-D**), whereas the XGB algorithm exhibited a steep increase (**Figure 2E**). Ultimately, all models reached a plateau as additional biomarkers were incorporated. Using RF, SVM, and XGB algorithms, we established three molecular panels comprising 8, 10, and 6 genetic predictors, designated as N8-RF-PCNSL (*MYD88*, *PIM1*, *CD79B*, *GNA13*, *IRF4*, *DTX1*, *KMT2D*, *B2M*), N10-SVM-PCNSL (*MYD88*, *PIM1*, *GNA13*, *SOCS1*, *CD58*, *EP300*, *PRDM1*, *B2M*, *BCL2*, *KMT2D*), and N6-XGB-PCNSL (*MYD88*, *PIM1*, *TP53*, *CD79B*, *NF1*, *IRF4*), respectively.

To evaluate the robustness and generalizability of the TIS-derived diagnostic panels, we conducted external validation and probabilistic calibration. In the TIS-based external validation set, N8-RF-PCNSL, N10-SVM-PCNSL and N6-XGB-PCNSL all achieved high diagnostic performance in terms of AUC, sensitivity, and specificity (**Figure 2F-H**). Log-loss and Brier score metrics in each of three panels indicated strong probabilistic calibration in both the TIS-based training and external validation sets (**Figure S4A-B**). Finally, no significant difference in diagnostic performance was observed between TIS-based training and external validation sets (**Figure S4C**).

CSF as the preferred liquid-biopsy source based on paired-sample concordance analysis

To establish the rationale for CSF-based panel validation, we performed mutational concordance analysis between TIS and liquid biopsy samples, quantified using Cohen's kappa coefficients and supported by sensitivity analyses. Across the 19 DMGs and ML-selected genetic biomarkers,

concordance of TIS-CSF was near-perfect, with Cohen's Kappa coefficients of 0.86, 0.89, 0.87, and 0.89, respectively. The agreement observed in each subgroup was significantly higher in CSF-TIS pairs than in PLA-TIS pairs ($P < 0.001$, **Figure 3A**). These findings were corroborated by sensitivity analyses (**Figure 3A** and **Figure S5A-B**), suggesting that CSF is a more suitable tissue-proximal liquid source for PCNSL diagnosis.

Diagnostic performance across CSF-based validations supported N8-RF-PCNSL as the preferred candidate panel

We first assessed CSF ctDNA detection rates using targeted sequencing data from the 112-gene panel applied in the BJTTH and BJFTH cohorts and the MSKCC-specific panels used in the publicly available MSKCC dataset. Across three CSF-based validations, CSF ctDNA positivity rates were high among patients with PCNSL (BJTTH: 86.3%; BJFTH: 92.6%; MSKCC: 60.9%). However, detection rates varied in non-PCNSL cases, including exclusively grade 4 glioma patients in the MSKCC cohort (BJTTH: 40.0%; BJFTH: 42.9%; MSKCC: 85.7%) (**Figure S6**). These findings indicate that mutation detection alone lacks sufficient diagnostic specificity, underscoring the need for classifier-based interpretation to achieve clinically actionable discrimination.

Then, we applied the three TIS-based ML-generated panels to CSF ctDNA data to evaluate their diagnostic performance for molecular identification of PCNSL. In the CSF-based BJTTH validation cohort, the N8-RF-PCNSL exhibited optimal diagnostic performance among the three ML-generated panels for distinguishing patients with PCNSL from those without, yielding an AUC of 0.90, a sensitivity of 79%, and a specificity of 100%. Although N10-SVM-PCNSL achieved slightly higher sensitivity, its specificity was lower (AUC: 0.88, sensitivity: 82%, specificity: 93%), whereas N6-XGB-PCNSL maintained 100% specificity but showed a lower sensitivity and AUC (AUC: 0.88, sensitivity: 77%, specificity: 100%; **Figure 3B**). Similar trends were observed in the CSF-based BJFTH validation cohort, where N8-RF-PCNSL and N6-XGB-PCNSL achieved an AUC of 0.94 while maintaining 100% specificity (**Figure 3C**). By contrast, the performance of all three panels was substantially reduced in the CSF-based MSKCC validation set. N8-RF-PCNSL and N6-XGB-PCNSL both showed higher specificity than N10-SVM-PCNSL; however, N10-SVM-PCNSL demonstrated the highest AUC and sensitivity (AUC: 0.69, sensitivity: 0.52), followed by N8-RF-PCNSL (AUC: 0.66, sensitivity: 0.43) and

N6-XGB-PCNSL (AUC: 0.62, sensitivity: 0.35) (**Figure S7A**). Although the diagnostic performance of the three panels was generally comparable, N8-RF-PCNSL showed a modest advantage over the other two panels in balancing discriminative performance, sensitivity, and specificity for identifying PCNSL.

Furthermore, we compared the diagnostic performance of the three ML-generated genetic biomarker panels with that of a single-mutation *MYD88* assay. In the CSF-based BJTTH validation cohort, all three panels exhibited higher sensitivity than the *MYD88* mutation assay. N8-RF-PCNSL and N6-XGB-PCNSL, unlike N10-SVM-PCNSL, maintained 100% specificity for PCNSL identification. Notably, N8-RF-PCNSL and N6-XGB-PCNSL significantly outperformed the *MYD88* mutation assay in terms of AUCs (0.90 vs. 0.84, $P = 0.003$; 0.88 vs. 0.84, $P = 0.01$; **Figure 3D**). Consistent findings were observed in the CSF-based BJFTH validation cohort. In the CSF-based MSKCC validation set, N8-RF-PCNSL and N10-SVM-PCNSL, but not N6-XGB-PCNSL, achieved higher AUC than the single-mutation *MYD88* assay (**Figure 3D and Figure S7B**). These results indicated that N8-RF-PCNSL, among the three panels, provided relatively consistent incremental diagnostic value over the single-mutation *MYD88* assay.

Considering that a clinically relevant subset of patients with PCNSL lack the canonical *MYD88* mutation, we further assessed the diagnostic performance of the ML-generated genetic biomarker panels in identifying *MYD88*^{wt} PCNSL. Among pathologically confirmed PCNSL patients, *MYD88*^{wt} status was identified in 41.1% (30/73), 33.3% (9/27), and 69.6% (16/23) of patients in the BJTTH, BJFTH, and MSKCC cohorts, respectively. In the CSF-based BJTTH validation cohort, all three panels demonstrated sensitivity ranging from 26% to 43%, with AUC values between 0.63 and 0.68. Notably, N8-RF-PCNSL outperformed N6-XGB-PCNSL in terms of sensitivity and both achieved 100% specificity for identifying PCNSL (**Figure 3E**). In the CSF-based BJFTH validation cohort, N8-RF-PCNSL and N6-XGB-PCNSL also maintained 100% specificity for identifying PCNSL, and all three panels yielded AUCs of at least 0.80 and a sensitivity of 0.67 (**Figure 3F**). In the CSF-based MSKCC validation set, although N8-RF-PCNSL and N6-XGB-PCNSL exhibited higher specificity than N10-SVM-PCNSL, none of the three panels achieved 100% specificity, and their overall diagnostic performance was limited, with AUCs ranging from 0.49 to 0.60 (**Figure S7A**). These results again

indicated that N8-RF-PCNSL provided a favorable overall balance among discriminative ability, sensitivity, and specificity for identifying *MYD88*^{wt} PCNSL.

Taken together, N8-RF-PCNSL was selected as the final candidate panel because of its favorable overall balance among discriminative performance, sensitivity, and specificity. Its compact eight-gene design further supports its clinical feasibility and downstream application.

Combined assessment of N8-RF-PCNSL and IL-10 levels for PCNSL diagnosis

To improve sensitivity while preserving specificity, we evaluated the diagnostic performance of combining N8-RF-PCNSL with CSF IL-10 levels. In the CSF-based BJTTH validation cohort, the N8-RF-PCNSL outperformed IL-10 testing, *MYD88* L265P mutation assays, and their combination (CSF IL-10 & *MYD88* L265P mutation status) in distinguishing PCNSL from non-PCNSL. Notably, the combined application of N8-RF-PCNSL and IL-10, where positivity for either or both tests was considered molecular evidence of PCNSL, performed best among five strategies, achieving a sensitivity of 86%, a specificity of 100%, and an AUC of 0.93 (**Figure 4**). Comparable findings were observed in the CSF-based BJFTH validation cohort. In the *MYD88*^{wt} subgroup, combined testing (CSF IL-10 & N8-RF-PCNSL) also yielded superior sensitivity and AUCs compared to either test alone, with diagnostic approaches maintaining 100% specificity (**Figure 4**).

Preliminary clinical application of CSF-based molecular diagnosis using N8-RF-PCNSL assisted with IL-10 testing

In the biopsy-infeasible real-world cohort, 26/42 (61.9%) patients were molecularly diagnosed as PCNSL using the CSF-based minimally invasive molecular diagnostic approach (**Figure 5A** and **Table S1**). Sixteen patients were not identified, generally showing CSF IL-10 levels < 5 pg/mL and negative molecular findings on N8-RF-PCNSL (P27 to P42). Three of the 26 patients showed an atypical molecular profile (*MYD88*^{wt}): one with mutations in both *PIM1* and *CD79B* (P26); and two with mutations in *PIM1* only (P11 and P12). In five patients (P3, P4, P5, P16, and P17), CSF-based liquid biopsy results were discordant, with IL-10 levels < 5 pg/mL but positive molecular findings on N8-RF-PCNSL. Among the 26 patients, 25 patients initiated a chemo-free regimen comprising zanubrutinib, a Bruton's tyrosine kinase inhibitor (BTKi), combined with lenalidomide, whereas one (P23) died before

treatment initiation. At the 3-month radiographic evaluation, contrast-enhanced MRI demonstrated consistent lesion regression in 25 patients, with 21 achieving complete remission and 4 achieving partial remission (**Figure 5B**). Representative cases were presented in **Figure 5C**.

Webpage deployment tool

We developed a user-friendly online interface to provide preliminary diagnostic insights based on input genetic biomarkers or their combinations. The web application is publicly accessible at: <https://n8-rf-pcnsl.shinyapps.io/n8-rf-pcnsl/>.

Discussion

Previous studies on CSF ctDNA-based targeted panels for PCNSL diagnosis are limited, and no consensus biomarker set has yet been established that optimally balances diagnostic accuracy, technical cost, and clinical deployability. To address this gap, our study introduced a genetics-based molecular panel, termed N8-RF-PCNSL, which integrates only eight highly specific genes. N8-RF-PCNSL demonstrated exceptional performance, achieving a sensitivity of at least 79% and a specificity of 100% in distinguishing patients with PCNSL from those without in CSF-based BJTTH and BJFTH validation cohorts. Collectively, these findings support a practical CSF-first, genetics-informed diagnostic strategy that may provide molecular evidence in support of PCNSL when tissue diagnosis is infeasible or delayed.

However, the performance of N8-RF-PCNSL was limited in the CSF-based MSKCC validation set, and several factors should be considered. First, heterogeneity of cross-platform sequencing may have contributed to the reduced diagnostic performance. Although all datasets were harmonized at gene-feature level using the 112-gene panel as reference before modeling, this strategy could not fully account for platform-specific differences in coding-region coverage, sequencing depth, analytical sensitivity, specimen processing, or variant-calling pipelines. In addition, we observed that the sequencing data from the CSF-based MSKCC validation set and the other two validation sets were not fully comparable, even after cross-platform gene-level harmonization (**Table S7**). Analysis for the MSKCC cohort should therefore be solely regarded as an exploratory CSF-based cross-platform

external validation of model rather than a definitive assessment of model generalizability. Second, population-related differences between the US and Chinese cohorts may also have contributed, given the potentially distinct genomic landscape of PCNSL reported in Chinese patients compared with non-Chinese populations [33]. Third, we performed a sample-level analysis to investigate why the model did not transfer effectively to the CSF-based MSKCC validation set. N8-RF-PCNSL misclassified 16 cases, including 13 false-negative PCNSL cases and 3 false-positive glioma cases. The false-negative results appeared to be primarily attributable to the lower mutation detection rate in the CSF-based MSKCC set than in the other two cohorts (**Figure S2**). This reduced cross-cohort comparability and likely weakened model performance. For the false-positive cases, the CSF samples harbored either *MYD88* or *PIM1* mutations, both of which are commonly associated with lymphoma and have rarely been reported in other intracranial neoplasms [13, 20]. However, the *MYD88* mutations were noncanonical G85D and L224F variants rather than the canonical L265P variant. The absence of variant-level mutation patterns from N8-RF-PCNSL may therefore have contributed to misclassification. In addition, all three CSF samples showed a high nonsynonymous mutation burden, and each contained at least one N8-RF-PCNSL feature-gene alteration at high variant allele frequency (VAF). However, the model incorporated neither VAF nor mutation burden, potentially limiting its ability to distinguish these cases accurately. Overall, future diagnostic model development should move beyond gene-level binarization by incorporating broader mutational features, while also relying on standardized pre-analytical procedures, sequencing platforms, and bioinformatic pipelines.

The EHA-ESMO guidelines recommend *MYD88* L265P mutation analysis and IL-10 measurement in CSF samples as supportive diagnostic strategies for PCNSL when brain biopsy is unfeasible [11, 12]. Our study demonstrates that the N8-RF-PCNSL assay exhibits greater sensitivity than either IL-10 measurement or *MYD88* L265P mutation testing, both individually and in combination, while maintaining a specificity of 100% across all methods. In the *MYD88*^{wt} subgroup, N8-RF-PCNSL exhibited sensitivity equivalent to that of IL-10 testing, with preserved specificity. Importantly, combining N8-RF-PCNSL with IL-10 measurement in CSF samples enhanced sensitivity in both overall and *MYD88*^{wt} populations, suggesting a synergistic diagnostic effect. Given the lack of standardized IL-10 cutoffs and the associated inter-laboratory variability, integrating N8-RF-PCNSL

may help circumvent underdiagnosis resulting from inconsistent IL-10 thresholds. The complementary roles of these assays warrant validation in larger, multicenter cohorts.

Compared to previously reported CSF ctDNA-based multigene models [19], N8-RF-PCNSL merely incorporates eight molecular markers, thus offering greater cost-efficiency and practical feasibility, particularly in resource-limited settings. To our knowledge, this work provides one of the first CSF ctDNA-based molecular panels to demonstrate high specificity for identifying *MYD88^{wt}* PCNSL (**Table S9**). An advantage of this multigene approach lies in its ability to identify a subset of PCNSL cases that would otherwise be missed by a single *MYD88* mutation assay. However, its sensitivity was limited and diagnostic performance varied across cohorts. Further validation and optimization are warranted.

In the biopsy-infeasible real-world cohort, given the activity of BTK inhibitors across B-cell lymphomas, treatment response alone cannot reliably substitute for biopsy as the diagnostic gold standard in most patients without definitive pathological confirmation. However, CSF-based liquid biopsy may provide a minimally invasive alternative to TIS biopsy under special clinical circumstances, particularly for frail patients, those receiving corticosteroids, or those with surgical contraindications (for example, long-term antiplatelet or anticoagulant use). Encouragingly, PCNSL cases were identified molecularly in over 60% of biopsy-infeasible patients by synergistic diagnostic strategy combining N8-RF-PCNSL with CSF IL-10 testing. Therefore, besides providing diagnostic support, this strategy is also promising to reduce diagnosis delays and strive for an optimal treatment opportunity for biopsy-infeasible patients with clinical deterioration, potentially improving their outcomes. Given the limited sample size and the absence of pathological confirmation, liquid-biopsy-guided therapy in biopsy-infeasible patients remains investigational. It should be regarded as a multidisciplinary salvage approach rather than an established standard-of-care substitute for histological diagnosis. In addition, several patients (P36, P37, and P39) with initially negative CSF-based liquid biopsy results were subsequently confirmed as PCNSL by TIS biopsy, indicating that a negative N8-RF-PCNSL/IL-10 result cannot exclude PCNSL when clinical suspicion remains high. Thus, the CSF-based liquid biopsy strategy should be regarded as a source of molecular evidence supporting a diagnosis of PCNSL rather than as a stand-alone diagnostic tool or a full replacement for

TIS biopsy. Discordant findings should prompt continued multidisciplinary reassessment and TIS confirmation whenever feasible.

Building on our initial experience with the combined N8-RF-PCNSL and CSF IL-10 testing diagnostic approach, we propose a preliminary diagnostic decision-making tree that provides guidance from minimally invasive to invasive workup in patients with radiologically suspected PCNSL (**Figure 6**). This proposed diagnostic workflow highlights the importance of CSF-based minimally invasive diagnostic strategy, facilitating stratified diagnostic management and timely access to treatment. We emphasize that this hypothesis regarding the decision-making tree is exploratory and requires prospective validation in a large, real-world patient population.

To facilitate clinical translation, we further characterized the distribution of mutational loci across the eight genes included in N8-RF-PCNSL (**Figure S8**). Consistent with previous studies [13, 34, 35], *MYD88* and *CD79B* harbored recurrent mutational hotspots, whereas the other six genes showed dispersed mutation patterns. Given that N8-RF-PCNSL is a compact eight-gene panel but predominantly comprises genes with dispersed mutation sites, an amplicon sequencing approach has the potential to capture the complete mutational spectrum of the panel. This approach may reduce not only testing costs but also turnaround time. Under an optimized workflow, testing from sample collection to result reporting can be completed within approximately 12 hours, facilitating timely diagnostic assessment and subsequent clinical management.

Several limitations must be acknowledged. First, the overall sample size remained relatively small, despite the multicohort design. Second, the spectrum of control diseases was limited. The Chinese non-PCNSL cohorts included few inflammatory or demyelinating lesions, brain metastases, or other non-neoplastic conditions, whereas the MSKCC cohort was restricted to grade 4 glioma.

Consequently, the specificity estimates from these cohorts cannot be generalized to the full spectrum of intracranial lesions that mimic PCNSL. Larger validation cohorts, particularly those adequately representing inflammatory and demyelinating lesions, are needed to establish the panel's specificity across clinically important differential diagnoses. Third, simple gene-level Boolean harmonization may be insufficient to ensure cross-platform comparability and may also mask mutation-specific information that contributes to diagnostic prediction, thereby reducing model performance. Fourth,

feature selection and panel-size determination were performed using the entire TIS-based training set without internal cross-validation, potentially yielding optimistic training-set AUC estimates. Although the panels were subsequently evaluated in an independent TIS-based external validation set, larger independent cohorts are needed to confirm their performance and assess the stability of gene selection. Overall, prospective validation using standardized sequencing and downstream analytical workflows is required before clinical implementation.

Conclusions

In summary, we developed N8-RF-PCNSL, a streamlined CSF ctDNA-based panel that provided molecular support for PCNSL diagnosis with high specificity in the two Chinese validation cohorts. Integration with CSF IL-10 further improved diagnostic sensitivity while maintaining specificity in these cohorts, suggesting a potential rule-in role when biopsy is infeasible or delayed. Prospective, multicenter studies incorporating standardized pre-analytical workflows and broader real-world differential diagnosis will be essential to confirm generalizability and refine clinical implementation.

Abbreviations

AUC: Area under the receiver operating characteristic curve, BJFTH: Beijing Fengtai Hospital, BJTTH: Beijing Tiantan Hospital, BTKi: Bruton's tyrosine kinase inhibitor, CGGA: Chinese Glioma Genome Atlas, CI: Confidence interval, CNS: Central nervous system, CR: Complete remission, CSF: Cerebrospinal fluid, ctDNA: Circulating tumor DNA, DLBCL: Diffuse large B-cell lymphoma, DMGs: Differentially mutated genes, ECOG: Eastern Cooperative Oncology Group, EHA-ESMO: European Hematology Association–European Society for Medical Oncology, FDR: False-discovery rate, FFPE: Formalin-fixed, paraffin-embedded, GBM: Glioblastoma, IL-10: Interleukin-10, IMiD: Immunomodulatory drug, Indels: Insertions/deletions, MAF: Mutant allele frequency, MDT: Multidisciplinary team, ML: Machine learning, MRI: Magnetic resonance imaging, MSKCC: Memorial Sloan Kettering Cancer Center, PCNSL: Primary central nervous system lymphoma, PCA: Principal component analysis, PLA: Plasma, PR: Partial remission, RF: Random forest, ROC: Receiver

operating characteristic, SNVs: Single-nucleotide variants, SVM: Support vector machine, TIS: Tumor tissue, VAF: variant allele frequency, WES: whole-exome sequencing, wt: Wild-type, XGB: XGBoost

Supplementary Material

Table S1-9

Figure S1-8

Acknowledgements

We sincerely thank all patients for their invaluable contributions to this study. We thank Yanhai Zhang from Burning Rock Biotech for assistance with sample collection, and Yin He from Burning Rock Biotech for assistance with data analysis. The authors disclose that no AI tools were used for data collection, analysis, or image manipulation; except for grammar and spelling review. The authors take full responsibility for the integrity of the manuscript.

Authors' contributions

Conceptualization: X. Ren, H. Li; Resources: X. Ren, S. Lin; Funding acquisition: X. Ren, Y. Mou, H. Duan; Project administration: X. Ren; Investigation: H. Li, X. Ren; Data curation: H. Li, M. Li, J. Zhang, G. Dong, L. Sang, L. Yuan, Q. Liu, L. Li, M. X. Li; Formal analysis: H. Li, X. Lv; Methodology: H. Li, X. Ren, H. Wang, L. Qin, Z. Zhao, B. Li, T. Zhu, X. Lv; Software: X. Lv, H. Li, H. Wang; Visualization: X. Lv, H. Li, H. Wang; Writing-original draft: H. Li, S. Shen; Writing-review and editing: X. Ren, Z. Zhao, H. Wang, L. Qin, H. Li, S. Shen, B. Li, T. Zhu, Z. Wang, D. Zhou, H. Duan, Y. Mou; Supervision: X. Ren, S. Lin, F. Meng, F. Li

Data and code availability

All data generated or analyzed from the Beijing Tiantan Hospital cohort and Beijing Fengtai Hospital cohort are available from the corresponding authors upon reasonable request (contact E-mail address: xiaohui ren@mail.ccmu.edu.cn). The data from the Memorial Sloan Kettering Cancer Center cohort utilized in this study are publicly accessible from online repositories (https://www.cbioportal.org/study/summary?id=pcnsl_msk_2024; http://www.cbioportal.org/study?id=glioma_msk_2018). The source code and complete prediction rules for N8-RF-PCNSL are publicly available at <https://github.com/PCNSLrxh/N8-model>.

Funding

This work was supported by the Capital's Funds for Health Improvement and Research to X. Ren (CFH 2026-1Q-1113); the National Natural Science Foundation of China to H. Duan (82573551) and Y. Mou (82373386); and the Key Science and Technology Project of Guangdong Province, China to Y. Mou (2023B0303020002).

Competing interests

No potential conflicts of interest were disclosed.

References

1. Ferreri AJM, Calimeri T, Cwynarski K, Dietrich J, Grommes C, Hoang-Xuan K, et al. Primary central nervous system lymphoma. *Nat Rev Dis Primers*. 2023; 9: 29.
2. Schaff LR, Grommes C. Primary central nervous system lymphoma. *Blood*. 2022; 140: 971-979.
3. Barajas RF, Politi LS, Anzalone N, Schoder H, Fox CP, Boxerman JL, et al. Consensus recommendations for MRI and PET imaging of primary central nervous system lymphoma: guideline statement from the International Primary CNS Lymphoma Collaborative Group (IPCG). *Neuro Oncol*. 2021; 23: 1056-1071.
4. Chen T, Liu Y, Wang Y, Chang Q, Wu J, Wang Z, et al. Evidence-based expert consensus on the management of primary central nervous system lymphoma in China. *J Hematol Oncol*. 2022; 15: 136.
5. Air EL, Leach JL, Warnick RE, McPherson CM. Comparing the risks of frameless stereotactic biopsy in eloquent and noneloquent regions of the brain: a retrospective review of 284 cases. *J Neurosurg*. 2009; 111: 820-824.
6. Ferreira MP, Ferreira NP, Pereira Filho Ade A, Pereira Filho Gde A, Franciscatto AC. Stereotactic computed tomography-guided brain biopsy: diagnostic yield based on a series of 170 patients. *Surg Neurol*. 2006; 65 Suppl 1: S27-S32.
7. Kellermann SG, Hamisch CA, Ruess D, Blau T, Goldbrunner R, Treuer H, et al. Stereotactic biopsy in elderly patients: risk assessment and impact on treatment decision. *J Neurooncol*. 2017; 134: 303-307.
8. Malone H, Yang J, Hershman DL, Wright JD, Bruce JN, Neugut AI. Complications Following Stereotactic Needle Biopsy of Intracranial Tumors. *World Neurosurg*. 2015; 84: 1084-1089.
9. Manoj N, Arivazhagan A, Bhat DI, Arvinda HR, Mahadevan A, Santosh V, et al. Stereotactic biopsy of brainstem lesions: Techniques, efficacy, safety, and disease variation between adults and children: A single institutional series and review. *J Neurosci Rural Pract*. 2014; 5: 32-39.
10. Bento LC, Correia RP, Alexandre AM, Nosawa ST, Pedro EC, Vaz ADC, et al. Detection of Central Nervous System Infiltration by Myeloid and Lymphoid Hematologic Neoplasms Using Flow Cytometry Analysis: Diagnostic Accuracy Study. *Front Med (Lausanne)*. 2018; 5: 70.
11. Ferreri AJM, Illerhaus G, Doorduijn JK, Auer DP, Bromberg JEC, Calimeri T, et al. Primary central

nervous system lymphomas: EHA-ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol.* 2024; 35: 491-507.

12. Ferreri AJM, Illerhaus G, Doorduijn JK, Auer DP, Bromberg JEC, Calimeri T, et al. Primary central nervous system lymphomas: EHA-ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Hemasphere.* 2024; 8: e89.

13. Calimeri T, Anzalone N, Cangi MG, Fiore P, Gagliardi F, Miserocchi E, et al. Molecular diagnosis of primary CNS lymphoma in 2024 using MYD88(Leu265Pro) and IL-10. *Lancet Haematol.* 2024; 11: e540-e549.

14. Ferreri AJM, Calimeri T, Lopedote P, Francaviglia I, Daverio R, Iacona C, et al. MYD88 L265P mutation and interleukin-10 detection in cerebrospinal fluid are highly specific discriminating markers in patients with primary central nervous system lymphoma: results from a prospective study. *Br J Haematol.* 2021; 193: 497-505.

15. Baraniskin A, Schroers R. Liquid Biopsy and Other Non-Invasive Diagnostic Measures in PCNSL. *Cancers (Basel).* 2021; 13: 2665.

16. Bravetti C, Degaud M, Armand M, Sourdeau E, Mokhtari K, Maloum K, et al. Combining MYD88 L265P mutation detection and clonality determination on CSF cellular and cell-free DNA improves diagnosis of primary CNS lymphoma. *Br J Haematol.* 2023; 201: 1088-1096.

17. Gupta M, Burns EJ, Georgantas NZ, Thierauf J, Nayyar N, Gordon A, et al. A rapid genotyping panel for detection of primary central nervous system lymphoma. *Blood.* 2021; 138: 382-386.

18. Heger JM, Mattlener J, Schneider J, Godel P, Sieg N, Ullrich F, et al. Entirely noninvasive outcome prediction in central nervous system lymphomas using circulating tumor DNA. *Blood.* 2024; 143: 522-534.

19. Mutter JA, Alig SK, Esfahani MS, Lauer EM, Mitschke J, Kurtz DM, et al. Circulating Tumor DNA Profiling for Detection, Risk Stratification, and Classification of Brain Lymphomas. *J Clin Oncol.* 2023; 41: 1684-1694.

20. Nayak L, Bettgowda C, Scherer F, Galldiks N, Ahluwalia M, Baraniskin A, et al. Liquid biopsy for improving diagnosis and monitoring of CNS lymphomas: A RANO review. *Neuro Oncol.* 2024; 26: 993-1011.

21. Olszewski AJ, Chorzalska AD, Petersen M, Ollila TA, Zayac A, Kurt H, et al. Detection of

- clonotypic DNA in the cerebrospinal fluid as a marker of central nervous system invasion in lymphoma. *Blood Adv.* 2021; 5: 5525-5535.
22. Wang Y, Springer S, Zhang M, McMahon KW, Kinde I, Dobbyn L, et al. Detection of tumor-derived DNA in cerebrospinal fluid of patients with primary tumors of the brain and spinal cord. *Proc Natl Acad Sci U S A.* 2015; 112: 9704-9709.
 23. Tivey A, Church M, Rothwell D, Dive C, Cook N. Circulating tumour DNA - looking beyond the blood. *Nat Rev Clin Oncol.* 2022; 19: 600-612.
 24. Hernandez-Verdin I, Kirasic E, Wienand K, Mokhtari K, Eimer S, Loiseau H, et al. Molecular and clinical diversity in primary central nervous system lymphoma. *Ann Oncol.* 2023; 34: 186-199.
 25. Radke J, Ishaque N, Koll R, Gu Z, Schumann E, Sieverling L, et al. The genomic and transcriptional landscape of primary central nervous system lymphoma. *Nat Commun.* 2022; 13: 2558.
 26. Grommes C, Nandakumar S, Schaff LR, Gavrilovic I, Kaley TJ, Nolan CP, et al. A Phase II Study Assessing Long-term Response to Ibrutinib Monotherapy in Recurrent or Refractory CNS Lymphoma. *Clin Cancer Res.* 2024; 30: 4005-4015.
 27. Miller AM, Shah RH, Pentsova EI, Pourmaleki M, Briggs S, Distefano N, et al. Tracking tumour evolution in glioma through liquid biopsies of cerebrospinal fluid. *Nature.* 2019; 565: 654-658.
 28. Jiang S, Qin Y, Jiang H, Liu B, Shi J, Meng F, et al. Molecular profiling of Chinese R-CHOP treated DLBCL patients: Identifying a high-risk subgroup. *Int J Cancer.* 2020; 147: 2611-2620.
 29. Qi F, Cao Z, Chen B, Chai Y, Lin J, Ye J, et al. Liquid biopsy in extranodal NK/T-cell lymphoma: a prospective analysis of cell-free DNA genotyping and monitoring. *Blood Adv.* 2021; 5: 2505-2514.
 30. Zhao Z, Zhang KN, Wang Q, Li G, Zeng F, Zhang Y, et al. Chinese Glioma Genome Atlas (CGGA): A Comprehensive Resource with Functional Genomic Data from Chinese Glioma Patients. *Genomics Proteomics Bioinformatics.* 2021; 19: 1-12.
 31. Cheng DT, Mitchell TN, Zehir A, Shah RH, Benayed R, Syed A, et al. Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology. *J Mol Diagn.* 2015; 17: 251-264.
 32. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of

metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med.* 2017; 23: 703-713.

33. Li S, Li D, Xia Z, Wu J, Ren J, Li Y, et al. Genomic landscape and molecular subtypes of primary central nervous system lymphoma. *Blood Cancer J.* 2026; 16: 8.

34. Nakamura T, Tateishi K, Niwa T, Matsushita Y, Tamura K, Kinoshita M, et al. Recurrent mutations of CD79B and MYD88 are the hallmark of primary central nervous system lymphomas. *Neuropathol Appl Neurobiol.* 2016; 42: 279-290.

35. Yamada S, Ishida Y, Matsuno A, Yamazaki K. Primary diffuse large B-cell lymphomas of central nervous system exhibit remarkably high prevalence of oncogenic MYD88 and CD79B mutations. *Leuk Lymphoma.* 2015; 56: 2141-2145.

1 Table 1. Clinical characteristics among BJTTH, BJFTH and MSKCC cohorts

	BJTTH cohort			BJFTH cohort		MSKCC cohort	
	PCNSL (N = 76)	Non-PCNSL Control* (N = 15)	Grade 4 Glioma (N = 54)	PCNSL (N = 27)	Non-PCNSL Control# (N = 14)	PCNSL (N = 29)	Grade 4 Glioma (N = 28)
Age, yrs, median (range)	63 (27-79)	49 (30-66)	54 (19-76)	68 (39-81)	53 (19-73)	69 (32-90)	-
Male, N (%)	43/76 (56.6)	10/15 (66.7)	26/54 (48.1)	16/27 (59.3)	11/14 (78.6)	18/29 (62.1)	19/28 (67.9)
ECOG, median (range)	2 (0-4)	1 (0-2)	-	2 (0-4)	0 (0-4)	1 (0-2)	-
Pre-biopsy CST, N (%)	3/76 (3.9)	0/15 (0)	-	1/27 (3.7)	0/14 (0)	-	-
Multiple lesions, N (%)	40/76 (52.6)	5/15 (33.3)	-	12/27 (44.4)	2/14 (14.3)	-	-
Lesions involving deep sites, N (%)	56/76 (73.7)	13/15 (86.7)	-	15/27 (55.6)	6/14 (42.9)	-	-
Tumor burden, cm³, median (range)	16.34 (1.37-81.89)	11.49 (3.46-74.44)	-	14.99 (4.84-97.50)	15.82 (3.18-60.31)	-	-
CSF IL-10 > 5 pg/mL, N (%)	31/76 (40.8)	0/15 (0)	-	22/27 (81.5)	0/14 (0)	-	-
Cytology positivity, N (%)	15/76 (19.7)	0/15 (0)	-	5/27 (18.5)	0/14 (0)	-	-
Hans, N (%)							
GCB	19/76	-	-	2/27	-	3/29	-

	(25.0)			(7.4)		(10.3)	
Non-GCB	45/76	-	-	25/27	-	21/29	-
	(59.2)			(92.6)		(72.4)	
Unknown	12/76	-	-	0/27	-	5/29	-
	(15.8)			(0)		(17.2)	

2 * Non-PCNSL brain lesions consist of glioblastoma (N = 4), IDH-mutant grade 4 astrocytoma (N = 1), anaplastic astrocytoma (N = 2), focal cerebral
3 inflammatory lesions (N = 2), demyelination (N = 2), brain metastasis (N = 1), gliosis (N = 1), cerebral infarction (N = 1) and craniopharyngioma (N = 1)

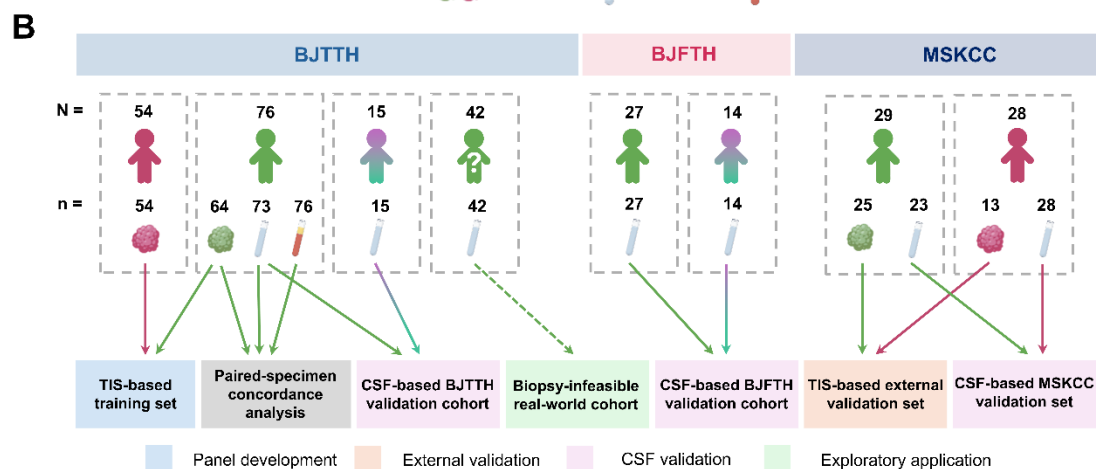
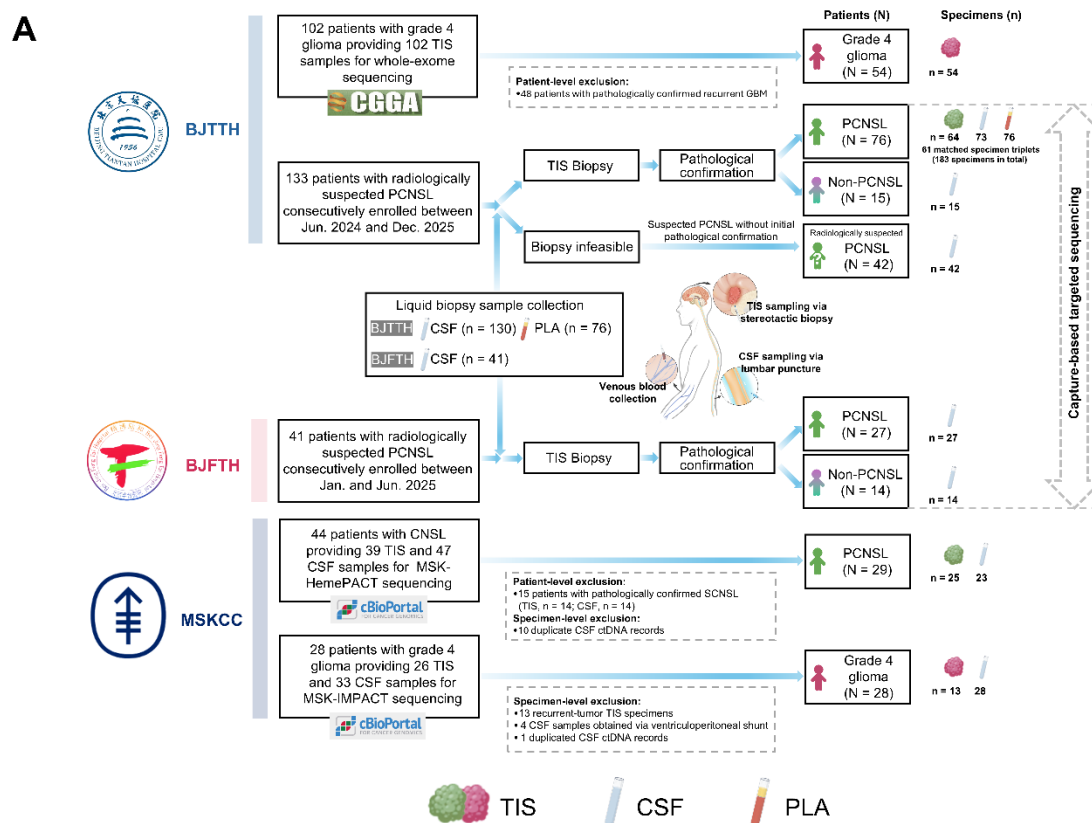
4 # Non-PCNSL brain lesions consist of glioblastoma (N = 10), IDH-mutant grade 4 astrocytoma (N = 1), demyelination (N = 2) and mesenchymal-origin
5 malignant brain tumor (N = 1)

6 BJFTH: Beijing Fengtai Hospital, BJTTH: Beijing Tiantan Hospital, CST: corticosteroid therapy, ECOG: Eastern Cooperative Oncology Group, GCB: germinal
7 center B-cell-like lymphoma, IL-10: interleukin-10, MSKCC: Memorial Sloan Kettering Cancer Center, N: number, PCNSL: primary central nervous system
8 lymphoma

9 Figure Legends

10 **Figure 1. Overview of the study population, specimens, and analytical roles.** (A) Flowchart
11 depicting patient enrollment, exclusion criteria, biofluid collection, patient-level diagnostic groups, and
12 the types and numbers of specimens available for each group across the BJTTH, BJFTH, and
13 MSKCC cohorts. Patients were classified according to pathological diagnosis, except for biopsy-
14 infeasible patients in the BJTTH cohort. These patients without initial pathological confirmation were
15 included because their lesions were radiologically suspected to be PCNSL. N denotes the number of
16 patients, whereas n denotes the number of specimens. Patient- and specimen-level exclusions are
17 shown in dashed boxes. (B) Allocation of patients and their corresponding specimens to the analytical
18 datasets. Colors at the bottom indicate the analytical roles of the respective datasets. Matched TIS,
19 CSF, and PLA specimens from the BJTTH cohort were used for paired-specimen mutational
20 concordance analysis. N denotes the number of patients, whereas n denotes the number of
21 specimens.

22 BJFTH: Beijing Fengtai Hospital, BJTTH: Beijing Tiantan Hospital, CNSL: central nervous system
23 lymphoma, CSF: cerebrospinal fluid, MSKCC: Memorial Sloan Kettering Cancer Center, PCNSL:
24 primary central nervous system lymphoma, PLA: plasma, SCNSL: secondary central nervous system
25 lymphoma, TIS: tumor tissue

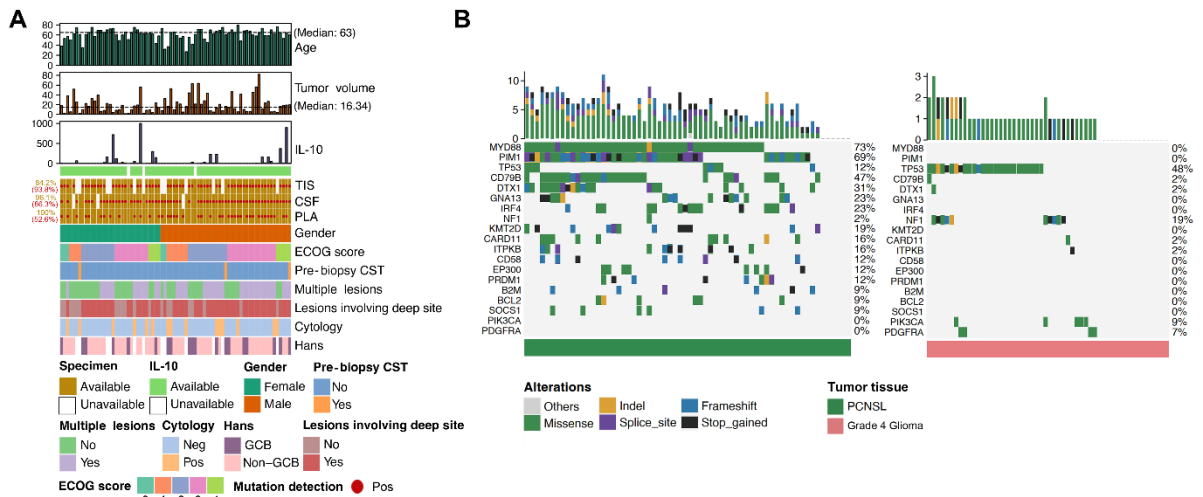


26

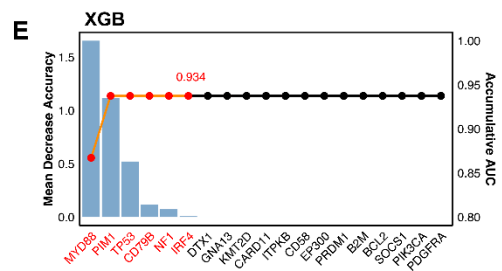
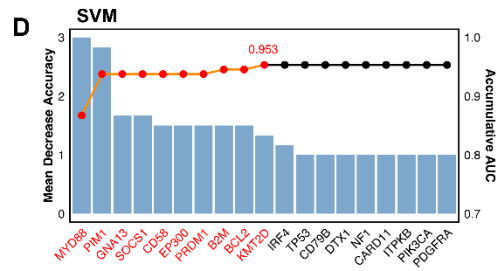
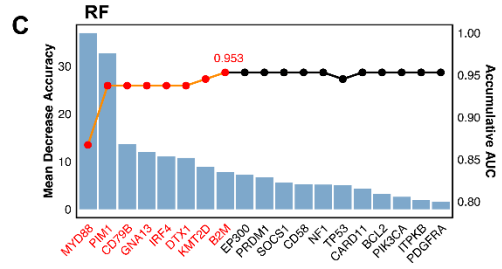
27

Figure 2. Clinical information and the development and validation of TIS-based genetic biomarker panels for diagnosing PCNSL. (A) Clinical baseline, pathological features and the specimen sources of patients with PCNSL from BJTTH cohort. (B) waterfall plots illustrating the mutation frequencies and genetic alterations of DMGs in PCNSL vs. grade 4 glioma samples. (C-E) ML-based sequential forward selection for identifying candidate genetic predictors and (F-H) assessing the diagnostic performance of the proposed biomarker panels in external validation. Bar charts show the ranked importance of individual mutated genes in the classification task (left axis). The line chart delineated cumulative AUCs (right axis) as each mutated gene is sequentially added. Candidate genetic predictors selected for final panel development are highlighted in red. Confusion matrices visualize model classification results, with color intensity reflecting the number of cases per category. ROC curves demonstrate the diagnostic performance of biomarker panels developed using RF, SVM, and XGB models, with corresponding AUCs.

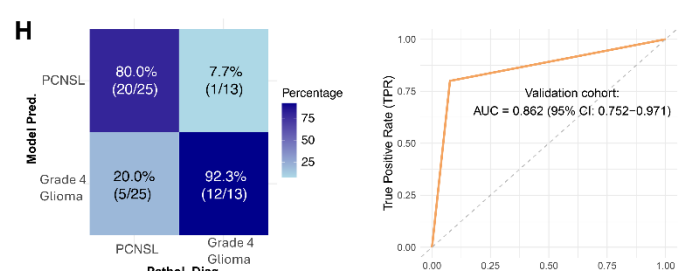
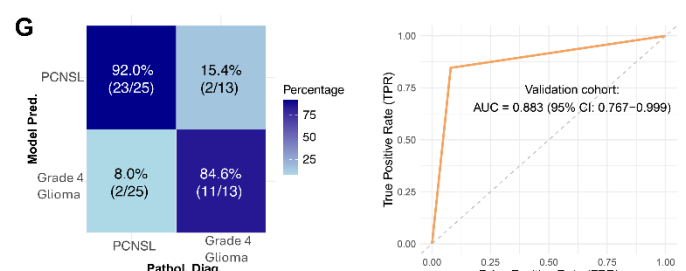
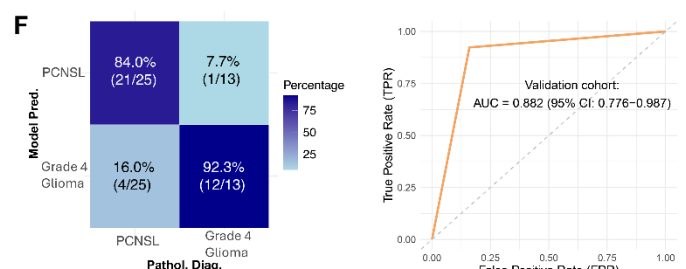
AUC: area under the receiver operating characteristic curve, CSF: cerebrospinal fluid, CST: corticosteroid therapy, DMGs: differentially mutated genes, ECOG: Eastern Cooperative Oncology Group, FPR: false positive rate, GCB: germinal center B-cell-like lymphoma, IL-10: interleukin-10, ML: machine learning, Neg: negative, Pathol. Diag.: pathological diagnosis, PLA: plasma, Pos: positive, Pred.: prediction, RF: random forest, ROC: receiver operating characteristic, SVM: support vector machine, TIS: tumor tissue, TPR: true positive rate, vs.: versus, XGB: XGboost



TIS-Based Biomarker Selection



TIS-Based External Validation



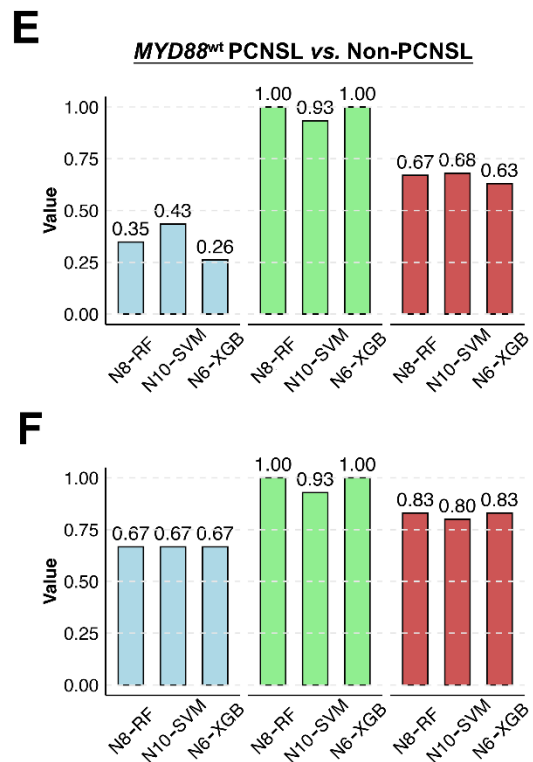
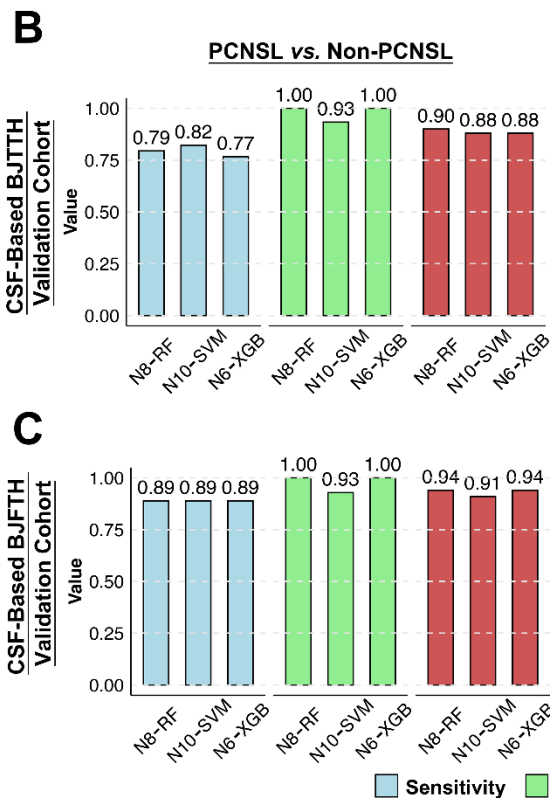
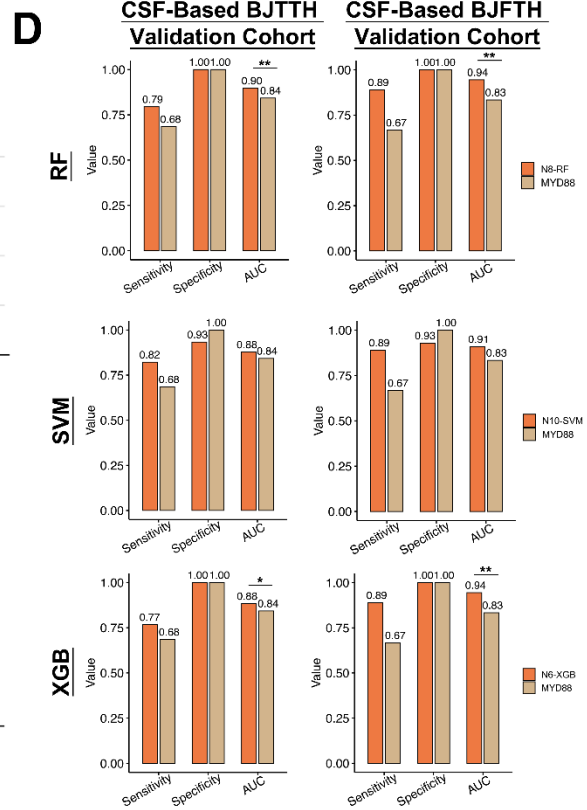
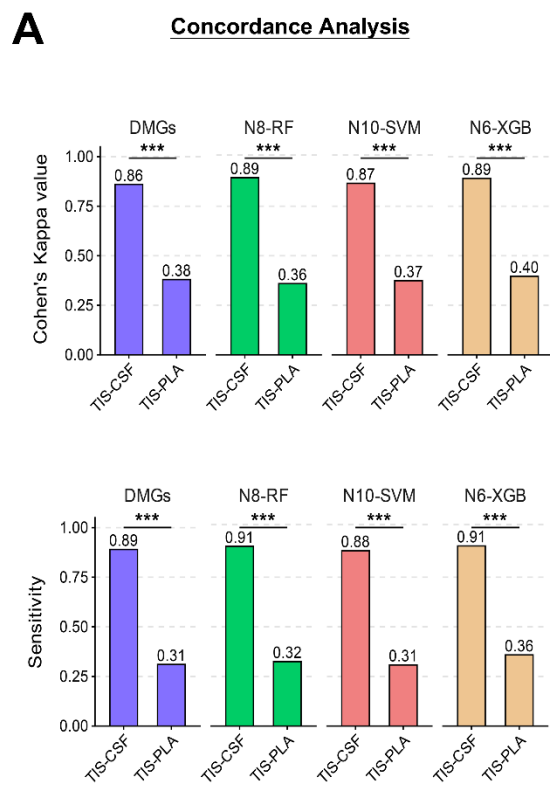
46

47

Figure 3. Mutational concordance analysis and diagnostic feasibility of the ML-generated panels in patients with PCNSL using CSF ctDNA. (A) Inter-sample agreement analysis of mutational profiles derived from TIS- and matched liquid specimen-informed sequencing. Cohen's Kappa coefficient (left axis) was calculated for DMGs and for mutated genes selected by ML algorithms. Sensitivity (left axis) was calculated based on mutated genes detected in TIS, which were either identified in CSF or PLA. Sensitivity is presented for both DMGs and ML-selected mutated genes. Cohen's Kappa coefficient was interpreted as: nearly perfect (0.81-1.00), substantial (0.61-0.80), moderate (0.41-0.60), fair (0.21-0.40), slight (0.01-0.20) and none (≤ 0). (B-C) Diagnostic performance of ML-generated panels using CSF ctDNA in differentiating patients with PCNSL from non-PCNSL. Bar charts display sensitivity, specificity, and AUC for the three established panels. BJTTH: $AUC_{RF} = 0.90$ (95% CI: 0.85-0.94), $AUC_{SVM} = 0.88$ (95% CI: 0.80-0.96), $AUC_{XGB} = 0.88$ (95% CI: 0.84-0.93). BJFTH: $AUC_{RF} = 0.94$ (95% CI: 0.88-1.00), $AUC_{SVM} = 0.91$ (95% CI: 0.82-1.00), $AUC_{XGB} = 0.94$ (95% CI: 0.88-1.00). (D) Comparative diagnostic performance of the N8-RF-PCNSL, N10-SVM-PCNSL and N6-XGB-PCNSL panels versus *MYD88*^{mut} mono-assay, respectively, in differentiating patients with PCNSL from non-PCNSL through CSF ctDNA detection. (E-F) Diagnostic performance of the ML-generated panels using CSF ctDNA in differentiating patients with *MYD88*^{wt} PCNSL from non-PCNSL. Bar charts compare sensitivity, specificity, and AUC among the three panels. BJTTH (*MYD88*^{wt} subgroup): $AUC_{RF} = 0.67$ (95% CI: 0.57-0.77), $AUC_{SVM} = 0.68$ (95% CI: 0.56-0.81), $AUC_{XGB} = 0.63$ (95% CI: 0.54-0.72). BJFTH (*MYD88*^{wt} subgroup): $AUC_{RF} = 0.83$ (95% CI: 0.67-1.00), $AUC_{SVM} = 0.80$ (95% CI: 0.62-0.98), $AUC_{XGB} = 0.83$ (95% CI: 0.67-1.00).

AUC: area under the receiver operating characteristic curve, BJFTH: Beijing Fengtai Hospital, BJTTH: Beijing Tiantan Hospital, CI: confidence interval, CSF: cerebrospinal fluid, DMGs: differentially mutated genes, ML: machine learning, MSKCC: Memorial Sloan Kettering Cancer Center, mut: mutant-type, PCNSL: primary central nervous system lymphoma, PLA: plasma, RF: random forest, SVM: support vector machine, TIS: tumor tissue, vs.: versus, wt: wild-type, XGB: XGboost

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$



75

76

Figure 4. Diagnostic performance of IL-10 testing, *MYD88* L265P mutation status, N8-RF-PCNSL, and their combinations for distinguishing PCNSL from non-PCNSL using CSF samples. Bar plots display the sensitivity (blue), specificity (green), and AUC (orange) for each diagnostic strategy across CSF-based BJTTH and BJFTH validation cohorts. Performance metrics are shown for IL-10 testing alone, *MYD88* L265P mutation status, the combination of IL-10 and *MYD88* L265P, N8-RF-PCNSL alone, and the combination of IL-10 with N8-RF-PCNSL, evaluated in both the overall population (PCNSL vs non-PCNSL) and the *MYD88*^{wt} subgroup (*MYD88*^{wt} PCNSL vs non-PCNSL).

AUC: area under the receiver operating characteristic curve, BJFTH: Beijing Fengtai Hospital, BJTTH: Beijing Tiantan Hospital, CSF: cerebrospinal fluid, IL-10: interleukin-10, RF: random forest classifier, PCNSL: primary central nervous system lymphoma

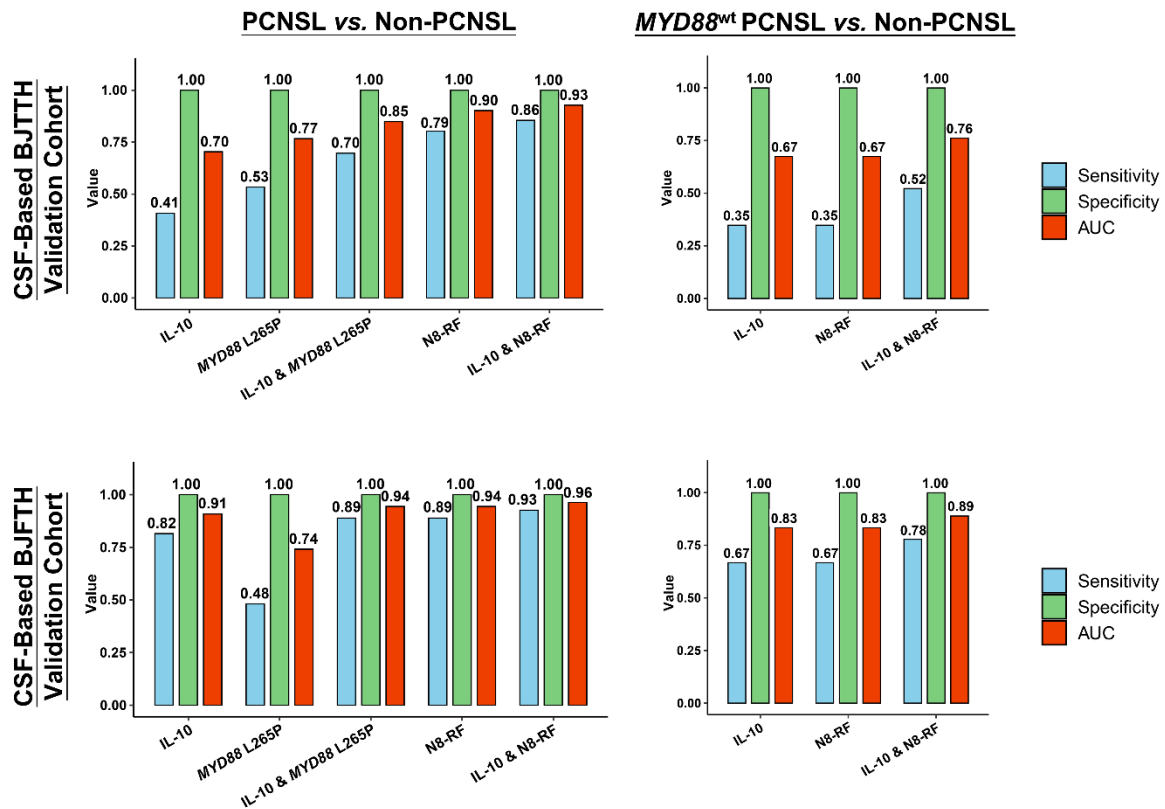
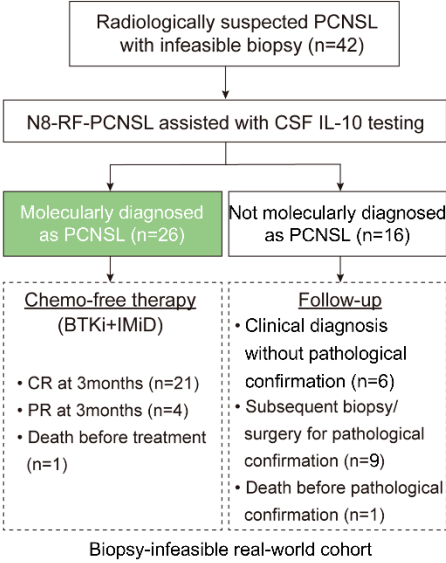


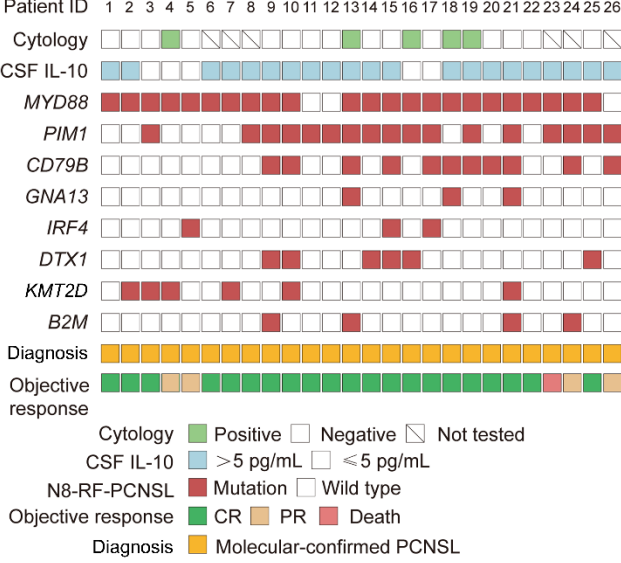
Figure 5. Molecular diagnosis and clinical outcomes in radiologically suspected, biopsy-infeasible PCNSL. (A) Clinical workflow of molecular-assisted diagnosis in biopsy-infeasible real-world cohort. (B) Individual patient-level molecular profiling and clinical response summary. Each column represents one patient. Gene mutation status in CSF tested by N8-RF-PCNSL, cytology results, CSF IL-10 levels, molecular diagnosis results, and treatment objective response are shown. Elevated IL-10 was defined as > 5 pg/mL. Treatment objective responses included CR, PR and death. Molecular diagnosis was primarily determined by the N8-RF-PCNSL testing. (C) Representative case (Case P16). The patient presented with a lesion involving the cingulate gyrus and corpus callosum, radiographically suggestive of GBM, PCNSL, or metastasis. Biopsy was considered infeasible due to high anesthesia risk and ongoing antiplatelet therapy. CSF analysis revealed positive molecular findings on N8-RF-PCNSL and positive cytology despite low IL-10 levels. The patient was molecularly diagnosed as PCNSL and subsequently received chemo-free therapy. Radiographic complete remission was observed at 3-month follow-up.

BTKi: Bruton tyrosine kinase inhibitor, CR: complete remission, CSF: cerebrospinal fluid, GBM: glioblastoma, IL-10: interleukin-10, IMiD: immunomodulatory drug, PCNSL: primary central nervous system lymphoma, PR: partial remission

A



B



C

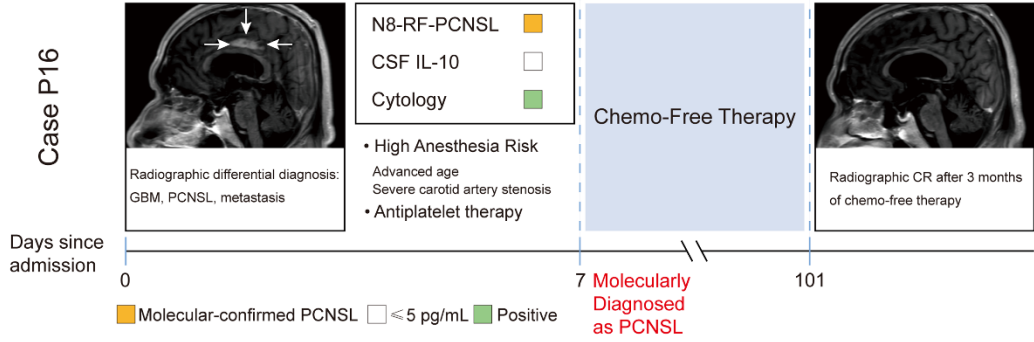
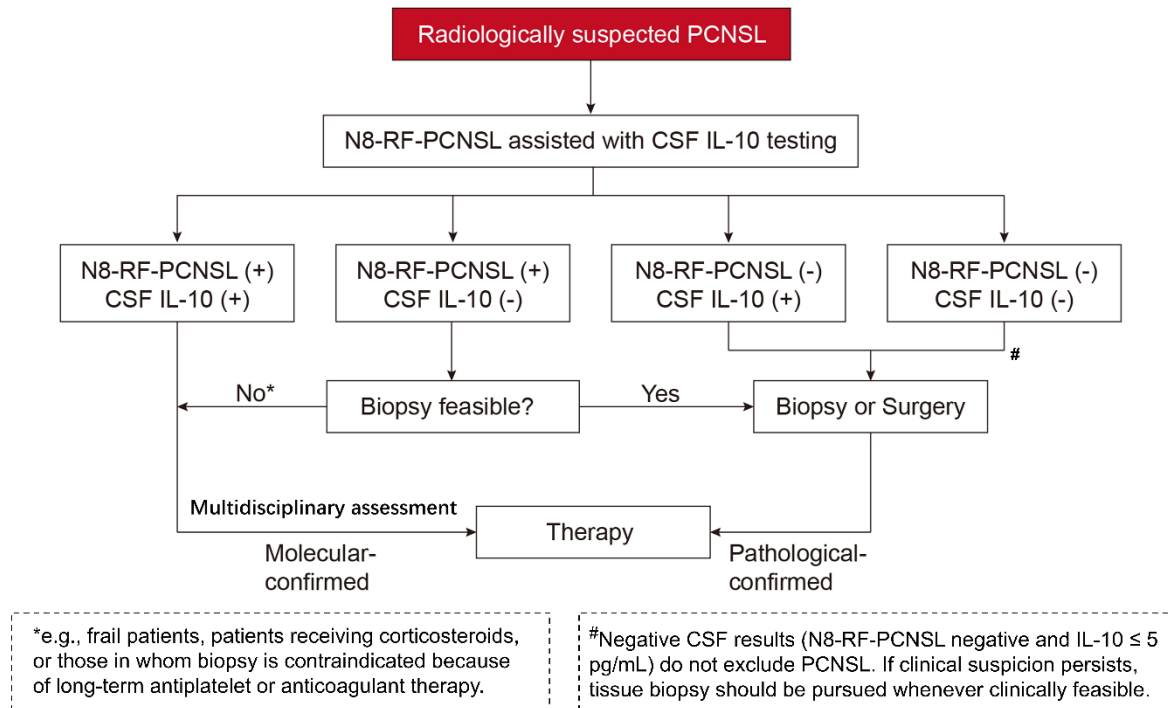


Figure 6. Diagnostic flow hypothesis focusing on CSF-based liquid biopsy. In this proposed diagnostic decision-making tree, patients with imaging suggestive of PCNSL first undergo CSF-based liquid biopsy using the N8-RF-PCNSL assisted with CSF IL-10 testing approach. Subsequent management is stratified based on the following results: (1) Patients who are classified as having PCNSL by N8-RF-PCNSL [i.e., N8-RF-PCNSL (+)] and have elevated CSF IL-10 levels may be molecularly diagnosed with PCNSL and proceed directly to treatment after multidisciplinary assessment. (2) For patients with N8-RF-PCNSL (+) but low CSF IL-10 levels, the next step is guided by clinical status: those with surgical contraindications, high anesthetic risk, or refusal of invasive procedures may also proceed to treatment after undergoing multidisciplinary assessment, while clinically stable patients should be considered for TIS biopsy to confirm histopathology. (3) Given that elevated CSF IL-10 levels are not entirely specific to PCNSL, the N8-RF-PCNSL prediction takes priority in this diagnostic hypothesis: patients who had N8-RF-PCNSL (-), even with elevated CSF IL-10 levels, should undergo invasive biopsy or surgery to establish a definitive pathological diagnosis before treatment. “+” and “-” denote the presence/absence of molecular evidence for diagnosing PCNSL by the N8-RF-PCNSL. The asterisk (*) denotes selected patients in whom biopsy is infeasible, particularly those who are frail, receiving corticosteroids, or have surgical contraindications, including those requiring long-term antiplatelet or anticoagulant therapy that cannot be safely interrupted. The hash symbol (#) denotes that negative CSF results (N8-RF-PCNSL negative and IL-10 $\leq$ 5 pg/mL) do not exclude PCNSL. If clinical suspicion persists, TIS biopsy should be pursued whenever clinically feasible.

CSF: cerebrospinal fluid, IL-10: interleukin-10, PCNSL: primary central nervous system lymphoma, TIS: tumor tissue



130

131