

ORIGINAL ARTICLE

Extranodal diffuse large B-cell lymphoma: Clinical and molecular insights with survival outcomes from the multicenter EXPECT study

Si-Yuan Chen¹ | Peng-Peng Xu¹ | Ru Feng² | Guo-Hui Cui³ | Li Wang¹ |
 Shu Cheng¹ | Rong-Ji Mu⁴  | Hui-Lai Zhang⁵ | Xiao-Lei Wei² |
 Yong-Ping Song⁶ | Kai-Yang Ding⁷ | Li-Hua Dong⁸ | Zun-Min Zhu⁹ |
 Shen-Miao Yang¹⁰ | Xin Wang¹¹  | Ting-Bo Liu¹² | Jian-Da Hu¹³  |
 Xiao-Yun Zheng¹⁴ | Ou Bai¹⁵ | Jing-Yan Xu¹⁶ | Liang Huang¹⁷ | Wei Sang¹⁸  |
 Ke-Qian Shi¹⁹ | Fan Zhou²⁰ | Fei Li²¹  | Ai-Bin Liang²² | Hui Zhou²³ |
 Si-Guo Hao²⁴ | Hong-Hui Huang²⁵ | Bin Xu²⁶  | Wen-Bin Qian²⁷ | Cai-Xia Li²⁸ |
 Zhi-Ming Li²⁹ | Chong-Yang Wu³⁰ | Xiao-Bo Wang³¹ | Wen-Yu Shi³² |
 Shu-Ye Wang³³ | Yu-Yang Tian³⁴ | Xi Zhang³⁵ | Ke-Shu Zhou⁸ | Li-Juan Cui³⁶ |
 Hui Liu³⁷ | Huo Tan³⁸ | Qing Leng³⁹ | Dong-Lu Zhao⁴⁰ | Ting Niu⁴¹ |
 Wei-Li Zhao^{1,42} 

Correspondence

Wei-Li Zhao, Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics, National Research Center for Translational Medicine at Shanghai, Ruijin Hospital, 197 Rui Jin Er Road, Shanghai 200000, P. R. China.
 Email: zhao.weili@yahoo.com

Abstract

Background: Diffuse large B-cell lymphoma (DLBCL) is the most common subtype of aggressive non-Hodgkin's lymphoma with distinct clinical and molecular heterogeneity. DLBCL that arises in extranodal organs is particularly linked to poor prognosis. This study aimed to determine the clinical and molecular charac-

Abbreviations: BN2, *BCL6* translocations and *NOTCH2* mutations; BP, biological process; *BTG1*, BTG anti-proliferation factor 1; *BTG2*, BTG anti-proliferation factor 2; CC, cellular component; CI, confidence interval; CNS, central nervous system; *CREBBP*, CREB binding protein; CT, computed tomography; DEG, differentially expressed genes; DLBCL, diffuse large B-cell lymphoma; ECOG, Eastern Cooperative Oncology Group; ENI, extranodal involvement; EXPECT, extranodal Project Collaborative Network; EZB, *EZH2* mutations and *BCL2* translocations; GATK, Genome Analysis Toolkit; GCB, germinal center B-cell; GI, gastrointestinal tract; GO, Gene Ontology; GSEA, Gene Set Enrichment Analysis; IPI, International Prognostic Index; KEGG, Kyoto Encyclopedia of genes and genomes; LDH, lactate dehydrogenase; LME, lymphoma microenvironment; MCD, *MYD88*^{L265P} and *CD79B* mutations; MF, molecular function; *MPEGL1*, macrophage-expressed 1; *MYD88*, myeloid differentiation primary response 88; N1, *NOTCH1* mutations; NCCN-IPI, National Comprehensive Cancer Network International Prognostic Index; NOS, Not Otherwise Specified; OS, overall survival; PCNSL, primary central nervous system lymphoma; PET-CT, ¹⁸F-fluorodeoxyglucose positron emission tomography with computed tomography; PFS, progression-free survival; PMBCL, primary mediastinal large B-cell lymphoma; PSM, propensity score matching; R-CEOP, rituximab plus cyclophosphamide, etoposide, vincristine, and prednisone; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone; SNP, single nucleotide polymorphism; SPSS, Statistical Package for the Social Sciences; ST2, *SGKI* and *TET2* mutations; *TBLIX1*, *TBLIX/Y*-related 1; *TET2*, tet methylcytosine dioxygenase 2; *TNFAIP3*, TNF α induced protein 3; *TP53*, tumor protein p53.

Si-Yuan Chen, Peng-Peng Xu, Ru Feng and Guo-Hui Cui contributed equally to this manuscript.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2025 The Author(s). *Cancer Communications* published by John Wiley & Sons Australia, Ltd on behalf of Sun Yat-sen University Cancer Center.

Funding information

Shanghai Clinical Research Center for Cell Therapy, Grant/Award Number: 23J41900100; National Natural Science Foundation of China, Grant/Award Number: 82130004; Clinical Research Plan of Shanghai Hospital Development Center, Grant/Award Number: SHDC2020CR1032B; National Key Research and Development Program, Grant/Award Number: 2022YFC2502600; Multicenter Clinical Research Project by Shanghai Jiao Tong University School of Medicine, Grant/Award Number: DLY201601

teristics of extranodal involvement (ENI) in DLBCL and assess the actual survival status of the patients.

Methods: In this population-based cohort study, we investigated the clinical features of 5,023 patients newly diagnosed with DLBCL. Their clinical conditions, eligibility criteria, and sociodemographic details were recorded and analyzed. Gene panel sequencing was performed on 1,050 patients to discern molecular patterns according to ENI.

Results: The 2-year overall survival (OS) rate was 76.2% [95% confidence interval (CI), 74.0%-78.2%], and the 5-year OS rate was 67.9% (95% CI, 65.2%-70.4%). The primary treatment was immunochemotherapy with rituximab. Specific lymphoma involvement sites, especially the bones, bone marrow, and central nervous system, were identified as independent adverse prognostic factors. A high prevalence of non-germinal center B-cell (non-GCB) phenotype and myeloid differentiation primary response 88 (*MYD88*)/*CD79B* mutations were noted in lymphomas affecting the breasts, skin, uterus, and immune-privileged sites. Conversely, the thyroid and gastrointestinal tract showed a low occurrence of non-GCB phenotype. Remarkably, patients with multiple ENIs exhibited a high frequency of *MYD88*, tet methylcytosine dioxygenase 2 (*TET2*), CREB binding protein (*CREBBP*) mutations, increased *MYD88*^{L265P} and *CD79B* mutation (MCD)-like subtypes, and poor prognosis. Genetic subtype-guided immunochemotherapy showed good efficacy in subgroup analyses after propensity score matching with 5-year OS and progression-free survival rates of 85.0% (95% CI, 80.6%-89.5%) and 72.1% (95% CI, 67.3%-76.7%).

Conclusions: In the rituximab era, this large-scale retrospective analysis from Asia confirmed the poor prognosis of DLBCL with multiple ENIs and underscored the efficacy of genetic subtype-guided immunochemotherapy in treating extranodal DLBCL.

KEYWORDS

Diffuse large B-cell lymphoma, disease progression, extranodal involvement, oncogenic mutation, prognosis, targeted therapy

1 | BACKGROUND

Diffuse large B-cell lymphoma (DLBCL) represents the most prevalent histological subtype of aggressive non-Hodgkin's lymphoma, exhibiting considerable clinical and molecular heterogeneity [1, 2]. Approximately one-third of DLBCL cases originate from extranodal organs, with the gastrointestinal (GI) tract, thyroid, testis, breasts, and skin being common sites [3–5], while DLBCL may also emerge in lymph nodes and subsequently disseminate to extranodal organs, including the bone marrow, lungs, pleura, liver, and central nervous system (CNS) [3, 6, 7]. Due to its clinical and prognostic heterogeneity, certain site-specific origins of extranodal DLBCL are classified as distinct pathologic subtypes, such as primary mediastinal large

B-cell lymphoma (PMBCL) [8], primary central nervous system lymphoma (PCNSL) [9], and primary cutaneous DLBCL, leg type [10].

The rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) regimen is the standard first-line treatment for most cases of extranodal DLBCL [11, 12]. Moreover, targeted agents have been progressively evaluated in clinical trials [13, 14]. The overall prognosis for extranodal DLBCL tends to be poorer than classical nodal DLBCL without extranodal involvement (ENI), although clinical outcomes vary significantly by site [6]. A large-scale analysis revealed that extranodal DLBCL was associated with a poor prognosis, with 5-year overall survival (OS) of 55.0% [15]. Another long-term study demonstrated a substantial improvement in 5-year OS for

extranodal DLBCL from 44.15% in the 1970s to 63.7% in the 2010s [16]. The number of involved extranodal sites served as a crucial prognostic indicator according to the International Prognostic Index (IPI) [17]. Additionally, the National Comprehensive Cancer Network International Prognostic Index (NCCN-IPI) identified the involvement of CNS, liver/GI tract, lungs, and bone marrow as adverse prognostic factors [18]. This disparity in prognosis has spurred broader research into extranodal variants beyond classical nodal DLBCL.

However, the limited genetic profiling studies for extranodal DLBCL have considerably hindered the interpretation of insights and their translation into clinical practice. Although retrospective studies in recent years have described the characteristics and prognosis of extranodal DLBCL [5, 6, 19, 20], these studies focused more on comparing patients with and without ENI, rather than systematically analyzing differences between them. In the present study, we initiated a multicenter retrospective study in China to elucidate the clinical and molecular characteristics, and to evaluate the therapeutic responses and prognostic factors of extranodal DLBCL.

2 | METHODS

2.1 | Study design and participants

This investigator-initiated study was a multicenter, retrospective observational trial. Eligible patients were ≥ 18 years; were pathologically diagnosed with DLBCL according to the 2016 World Health Organization classification standards, including non-specific types and various special subtypes; were newly diagnosed with DLBCL with involvement of extranodal sites; received clinical treatment for lymphoma; had measurable lesions, and underwent at least one valid efficacy assessment. Exclusion criteria were receiving only supportive care or not having access to valid efficacy assessment data. Thirty-nine Chinese hospitals affiliated with the Extranodal Project Collaborative Network (EXPECT) contributed to the research (NCT06549361). This study encompassed a dataset of 5,023 patients newly diagnosed with DLBCL between September 2002 and December 2022, as recorded in the registry data. All enrolled participants presented definite ENI. Nodal tissues and organs were defined to include lymph nodes, spleen, thymus, and Waldeyer's ring. Data collection was consistent with the STROBE guideline for observational studies and was approved by the Institutional Review Boards of all participating centers, adhering to the Declaration of Helsinki [21], with informed consent obtained from all subjects.

2.2 | Patients follow-up

Treatment response was evaluated at the end of treatment, based on a ^{18}F -fluorodeoxyglucose positron emission tomography with computed tomography (PET-CT) scan according to the Lugano criteria for non-Hodgkin's lymphoma, using the Deauville 5-point scale. Deauville score of 5 was considered progressive disease [22]. Follow-up examination was repeated every 3 months during the first year, 6 months during the second year, and every subsequent year. Follow-up examinations included a clinical examination, laboratory analysis, and contrast-enhanced neck, thorax, abdomen and pelvis computed tomography (CT). All follow-up information is available in the electronic data systems of the participating centers, with the last follow-up conducted on July 1, 2023.

2.3 | Data collection

The dataset included detailed patient information at the time of diagnosis, treatment approaches and responses to each treatment line, dates of diagnosis, relapse or progression, and dates of death or the last follow-up. Covariates for clinical characteristics and subgroup analyses within the dataset encompassed age, sex, Eastern Cooperative Oncology Group (ECOG) performance status, Ann Arbor stage, IPI score, treatment modalities, sites of ENI, and status of ENI (primary and secondary). Data were systematically extracted from electronic medical records by a team of research coordinators trained in medical chart review and adherence to a standardized data collection protocol. A central electronic database was established for online data collection. Monthly data quality control was conducted by a dedicated data surveillance team, which also provided feedback to all participating centers.

2.4 | Study outcomes

The primary objective of this study was to evaluate the prognosis and efficacy of treatment for extranodal DLBCL and explore the optimal treatment strategies in real-world populations. The primary outcome measures were OS and progression-free survival (PFS). The secondary outcome measures were the clinical and biological data filled into registry forms by physicians and data managers, including lactate dehydrogenase levels, clinical stage, performance status, age, and the number and location of extranodal sites.

2.5 | Gene panel sequencing

A total of 1,050 patients had enough tumor tissue samples remained after pathologic examination. Targeted sequencing of the lymphoma-related genes was performed on the 1,050 formalin-fixed and paraffin-embedded tumor samples using MultipSeq Custom Panel (Shanghai Rightongene Biotechnology Co., Ltd, Shanghai, China). All tumor tissues were obtained via core needle biopsy of lymphoma lesions, including both nodal and extranodal sites. The capture probes were designed based on ~0.39 Mb genomic regions of 55 lymphoma-related genes that are frequently mutated in DLBCL, other common lymphoma and hematologic malignancies [23]. In summary, gDNA was fragmented, end repaired, linked with sequencing adapters, purified and went through pre-PCR using the Enzyme Plus Library Prep Kit (iGeneTech, Beijing, China). After hybridization and concentration, sequencing was performed on the Novaseq (<http://www.illumina.com/>) sequencing platform. After sequencing, the quality of the raw sequencing data was assessed using FastQC software (version 1.11.4, <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). In addition, Trimmomatic software (version 3.6, <https://github.com/usadellab/Trimmomatic>) was used to process raw sequencing data to remove adaptor sequences and low-quality fragments. The original sequencing was aligned with the human reference genome hg19 using Burrows-Wheeler Aligner (0.7.13-r1126). The repeats were eliminated, and the base quality was recalibrated. Genome Analysis Toolkit (GATK) of Broad Institute (Cambridge, MA, USA) was used for single nucleotide polymorphism (SNP) calling. Mutation detection and analysis of Binary Alignment Map files were performed using the cancer genome analysis program Mutect2 in GATK at the Broad Institute and annotated with ANNOVAR (<https://annovar.openbioinformatics.org/en/latest/>). Also, FIR2 and F2R1 annotations were adopted, and FilterByOrientationBias was performed to filter the orientation bias. The single nucleotide variants (SNVs) and insertions and deletions (Indels) were screened based on the filtering conditions: (1) variants with mapping quality > 30 were retained; (2) SNVs or Indels with a mutation allele frequency < 0.001 in databases of 1000 genomes project, 1000 genome East Asian, ExAC all or ExAC East Asian and genomAD were retained; (2) SNVs or Indels with a VAF \geq 5% was retained; (3) database of SNP (v147) sites existed in the COSMIC database (<https://cancer.sanger.ac.uk/cosmic>) were retained; (4) SNPs or Indels including stopgain, stoploss, frameshift, nonframeshift and splicing sites were retained; (5) missense mutations with sift \leq 0.05, Polyphen2_HVAR_pred \geq 0.447 and combined annotation-dependent depletion > 4 were retained.

2.6 | Definition of treatment options and prognostic analysis

R-CHOP was defined as a standard regimen using rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. Among 4,784 patients who received rituximab-based regimens as first-line therapy, 3,027 (63.3%) were treated with standard R-CHOP (Supplementary Table S1). The univariate and multivariate analyses were performed in the R-CHOP cohort. Variables significant in univariate analysis were included in the multivariate model to identify independent prognostic factors. The dose-adjusted R-CHOP regimen was defined as R-CHOP-based regimen and combined with the R-CHOP regimen for the survival analyses. Chemo-free regimens were defined as regimens in which no chemotherapeutic drugs were used, and only targeted agents such as rituximab and small molecule agents were used. Substituting etoposide for doxorubicin in the standard R-CHOP was defined as R-CEOP and typically recommended for patients with contraindications to anthracyclines. Therapeutic regimens based on high-dose methotrexate was used in patients with CNS involvement.

DLBCL subtypes were determined utilizing the LymphPlex probabilistic classification [24]. It was a simplified 38-gene algorithm based on the information on mutations of 35 genes and rearrangements of three genes (*BCL2*, *BCL6*, and *MYC*) using whole exome/genome sequencing, RNA-sequencing, and fluorescence in situ hybridization in 337 newly diagnosed DLBCL patients, identifying seven distinct genetic subtypes: tumor protein p53 (*TP53*) mutation, the L265P mutations of myeloid differentiation primary response 88 (*MYD88*) and *CD79B* mutations (MCD)-like, *BCL6* fusions and *NOTCH2* mutations (BN2)-like, *NOTCH1* mutations (N1)-like, enhancer of zeste homolog 2 (*EZH2*) mutations and *BCL2* translocations (EZB)-like, serum/glucocorticoid regulated kinase 1 (*SGKI*) and tet methylcytosine dioxygenase 2 (*TET2*) mutations (ST2)-like, and not otherwise specified (NOS). Sensitivity, specificity, and precision (positive predictive value) of the LymphPlex algorithm for the subtype assignments were comparable to the assignments using the LymphGen algorithm [22], and the LymphPlex subtypes generally belonged to the corresponding LymphGen subtypes [24, 25]. The GUIDANCE-01 trial had validated the feasibility of immunochemotherapy based on genetic subtypes of DLBCL [26]. Consequently, we aimed to stratify patients according to genetic subtypes and defined subgroups of R-CHOP-X. Within the framework of standard R-CHOP, patients meeting any of the following criteria were categorized as the R-CHOP-X group: N1-like and unclassifiable subtypes received lenalidomide; MCD-like

and BN2-like subtypes received ibrutinib; EZB-like subtypes received tucidinostat; or *TP53*^{Mut} subtypes received decitabine. Patients receiving targeted agents outside of the standard R-CHOP but not adhering to the above criteria were also classified into the R-CHOP-based regimen.

We performed a propensity score matching (PSM) to assess treatment effects. A logistic regression model including age, ECOG performance status, Ann Arbor stage, serum lactate dehydrogenase (LDH) level, and number of ENI was used to estimate the PSM, and two groups (the R-CHOP-X group and the R-CHOP/R-CHOP-based group) with comparable baselines and the same number of people were obtained ($n = 239$ for each group). PSM for selected variables used logistic regression with the nearest-neighbor method. Calliper settings were set to less than 0.2 to restrict distance between matched units. We used clustering to account for differences between patients with similar propensity scores. A standardized mean difference threshold of less than 0.1 was determined as an indicator of reduced bias between treatment cohorts.

2.7 | RNA sequencing

RNA sequencing was performed on qualified frozen DLBCL tissues from 434 patients who had enough tumor tissue samples remained after pathologic examination and gene panel sequencing. For cases with single ENI, the tumor samples were taken from this extranodal site, and for cases with multiple ENI, the tumor samples were taken from one of the extranodal sites. Total RNA was extracted from the frozen tumor tissue using Trizol and the RNeasy Mini Kit (Qiagen). RNA quantity was measured using Nanodrop, and the integrity of the total RNA was evaluated using the RNA 6000 Nano Kit on an Agilent 2100 Bioanalyzer. The read pairs were aligned to the Refseq hg19 reference genome using the Burrows-Wheeler Aligner version 0.7.13-r1126. HTSeq (<http://htseq.readthedocs.io/>) was used to tabulate Transcript counts, and visual inspection was employed to rule out potential false-positive results. Bioinformatics analyses were conducted using R version 3.5.1. The R package “sva” was used to mitigate batch effects. Raw reads were normalized, and differentially expressed genes (DEGs) were identified using the R package “limma” (version 3.38.3). Patients were categorized into four lymphoma microenvironment (LME) subtypes based on the analysis of 25 functional gene expression signatures and a k-nearest neighbors supervised model [27]. This model was trained using a cohort of 4,656 DLBCL patients (<https://github.com/bostongene/lme>). Gene expression signatures obtained from SignatureDB (<https://gdc.cancer.gov/about-data/publications/DLBCL-2018>) were quantified using the single-sample

Gene Set Enrichment Analysis (GSEA) method (<https://www.gsea-msigdb.org/gsea/>) and were compared across different genetic or LME subtypes. For patients with RNA-seq data, the cell of origin was determined based on the expression of signatures identified in the Lymph2Cx assay.

2.8 | Gene set enrichment analysis

GSEA v4.0.1 software (<https://www.gsea-msigdb.org/gsea/index.jsp>) and the Molecular Signature Database (<https://www.gsea-msigdb.org/gsea/msigdb/>) were used. GSEA results were exhibited as either the upregulation or downregulation of the targeted gene set. The default metric method, Signal2Noise, was used to rank the genes. The analysis underwent 1,000 permutations to evaluate the statistical significance of the enrichment score, adhering to the recommendations of the GSEA team (<http://www.broadinstitute.org/gsea>). Pathways were deemed statistically significant when the P value < 0.05 and the false discovery rate < 0.25 .

2.9 | Gene Ontology and Kyoto Encyclopedia of Genes and Genomes enrichment analysis

Gene Ontology (GO) (<http://geneontology.org/>) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (<https://www.kegg.jp/>) enrichment analysis were used. For functional enrichment analysis, all DEGs were mapped to terms in the GO databases, and then significantly enriched GO terms were searched among the DEGs using $P < 0.050$ as the threshold. All DEGs were mapped to the KEGG database, and searched for significantly enriched KEGG pathways.

2.10 | Tumor immunophenotyping

The anti-tumor immunity activity score was generated using tumor immunophenotyping (<http://biocc.hrbmu.edu.cn/TIP>) which enabled the differentiation of specific T-cell subset recruitment. This approach incorporated 178 signature genes and 23 signature gene sets involved in the cancer-immunity cycle.

2.11 | Statistical analysis

Patient baseline characteristics were assessed using either the Pearson's χ^2 test or Fisher's exact test. Variations in immunity activity scores and normalized gene expression

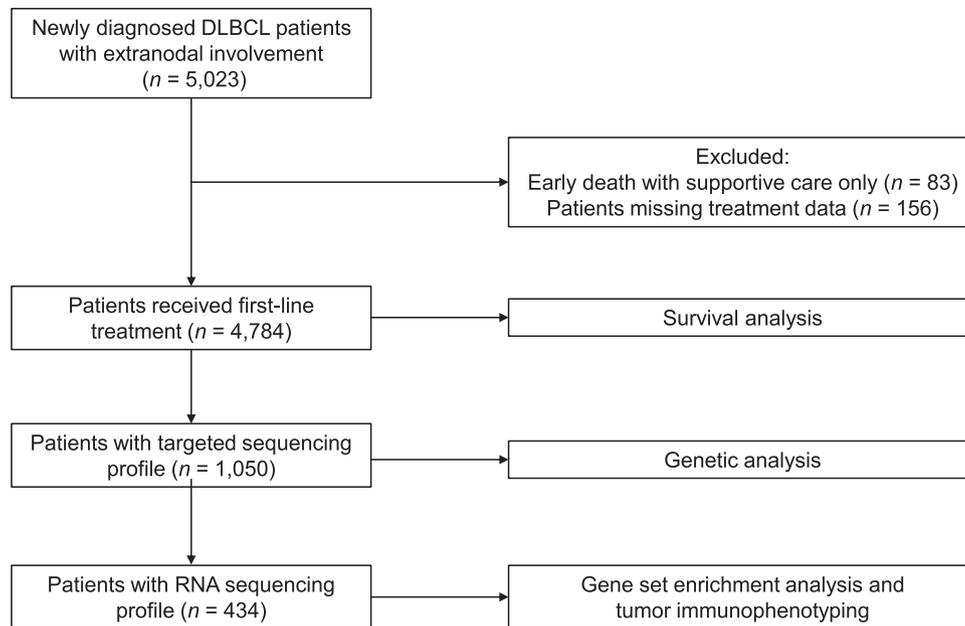


FIGURE 1 Flow chart of the patient selection and methodology. Abbreviations: DLBCL, diffuse large B-cell lymphoma.

between groups were examined using the Mann-Whitney *U* test. Correlations between mutations were calculated using Spearman's rank correlation coefficient. PFS was calculated from the date of diagnosis to either the date of disease progression/relapse detection or the date of the last follow-up. OS was calculated from the date of diagnosis to either the date of death or the date of the last follow-up. Survival curves were generated using the Kaplan-Meier method and compared using the log-rank test. Univariate hazard analysis was conducted using the Cox regression method. Significant variables identified in the univariate analysis were retained in the multivariate analysis. Statistical significance was defined as $P < 0.05$, and all reported P values in this manuscript were unadjusted. The statistical analyses described above were performed using Statistical Package for the Social Sciences (SPSS) version 26.0 software (SPSS Inc., Chicago, IL, USA).

3 | RESULTS

3.1 | Patient characteristics

A total of 4,784 patients newly diagnosed with DLBCL who received first-line treatment were included, comprising individuals with either single or multiple ENIs. The process of patient selection is depicted in Figure 1, and the clinical profiles of these patients are summarized in Table 1. Among all, 2,524 (52.8%) were male, and 2,260 (47.2%) were female. The median age at diagnosis was 62 years, with a range spanning from 18 to 101 years. A

majority of the patients presented with advanced stage III-IV disease (56.4%) and were categorized as high-risk according to IPI. Notably, multiple ENIs were significantly associated with a higher ECOG performance status ($P < 0.001$), age > 60 ($P < 0.001$), advanced Ann Arbor stage ($P < 0.001$), and elevated serum LDH levels ($P < 0.001$), compared to single ENI.

3.2 | Distribution of extranodal sites

Extranodal DLBCL was most commonly observed in the GI tract, including the stomach (18.4%) and intestine (14.9%). Additionally, the bones, bone marrow, CNS, lungs, kidney/adrenal glands, breasts, and liver accounted for more than 5%. In addition to these specific sites of ENI, there were cases in which the site of ENI was difficult to define, such as subcutaneous and intermuscular soft tissues, or atypical sites of ENI, such as the Waldeyer's ring and the spleen, which were not included in the site-specific analyses of this study (Table 1 and Figure 2A). Multiple ENIs were more frequently observed in organs such as the bones, bone marrow, kidney/adrenal glands, lungs, liver, pancreas, skin, nasal cavity, and uterus/ovaries, compared to single ENI (Figure 2B). Significant correlations were identified among ENI at various sites (Figure 2C). Simultaneous involvement of the bones and liver as well as simultaneous involvement of the kidney/adrenal glands and liver or pancreas were common. Involvements of the GI tract and CNS showed negative correlations with involvement of other extranodal sites.

TABLE 1 Baseline demographic and disease characteristics of patients with extranodal diffuse large B-cell lymphoma who received first-line treatment.

Characteristic	Whole cohort [cases (%)]	Number of extranodal involvement [cases (%)]		P value
		Single	Multiple	
Total	4,784	3,163	1,621	
Gender				<0.001
Male	2,524 (52.8)	1,612 (51.0)	912 (56.3)	
Female	2,260 (47.2)	1,551 (49.0)	709 (43.7)	
Age				<0.001
≤60 years	2,544 (53.2)	1,742 (55.1)	802 (49.5)	
>60 years	2,240 (46.8)	1,421 (44.9)	819 (50.5)	
ECOG score				<0.001
0-1	4,071 (85.1)	2,755 (87.1)	1,316 (81.2)	
≥2	713 (14.9)	408 (12.9)	305 (18.8)	
Ann Arbor stage				<0.001
I-II	2,087 (43.6)	1,886 (59.6)	201 (12.4)	
III-IV	2,697 (56.4)	1,277 (40.4)	1,420 (87.6)	
LDH				<0.001
Normal	2,306 (48.2)	1,801 (56.9)	505 (31.2)	
Elevated	2,478 (51.8)	1,362 (43.1)	1,116 (68.8)	
IPI				<0.001
0-2	2,886 (60.3)	2,459 (77.7)	427 (26.3)	
3-5	1,898 (39.7)	704 (22.3)	1,194 (73.7)	
Cell of origin (Hans)				0.023
GCB	1,345 (28.1)	907 (28.7)	438 (27.0)	
Non-GCB	2,561 (53.5)	1,710 (54.1)	851 (52.5)	
Not reported	878 (18.4)	546 (17.3)	332 (20.5)	
Double-hit/triple-hit lymphoma				<0.001
Yes	106 (2.2)	55 (1.7)	51 (3.1)	
No	1,782 (37.2)	1,136 (35.9)	646 (39.9)	
Not reported	2,896 (60.5)	1,972 (62.3)	924 (57.0)	
Extranodal sites				<0.001
Stomach	879 (18.4)	605 (19.1)	274 (16.9)	
Intestine	712 (14.9)	422 (13.3)	290 (17.9)	
Bones	642 (13.4)	186 (5.9)	456 (28.1)	
Bone marrow	547 (11.4)	142 (4.5)	405 (25.0)	
CNS	417 (8.7)	324 (10.2)	93 (5.7)	
Lungs	337 (7.0)	118 (3.7)	219 (13.5)	
Kidney/adrenal glands	336 (7.0)	89 (2.8)	247 (15.2)	
Breasts	318 (6.6)	231 (7.3)	87 (5.4)	
Liver	295 (6.2)	35 (1.1)	260 (16.0)	
Nasal cavity	220 (4.6)	102 (3.2)	118 (7.3)	
Mediastina	212 (4.4)	153 (4.8)	59 (3.6)	
Testis	210 (4.4)	125 (4.0)	85 (5.2)	
Pancreas	170 (3.6)	50 (1.6)	120 (7.4)	
Thyroid	162 (3.4)	87 (2.8)	75 (4.6)	
Skin	115 (2.4)	52 (1.6)	63 (3.9)	
Uterus/ovaries	107 (2.2)	31 (1.0)	76 (4.7)	
Others	928 (19.4)	411 (13.0)	517 (31.9)	

Abbreviations: ECOG, Eastern Cooperative Oncology Group; IPI, International Prognostic Index; GCB, germinal center B-cell; LDH, lactate dehydrogenase; CNS, central nervous system.

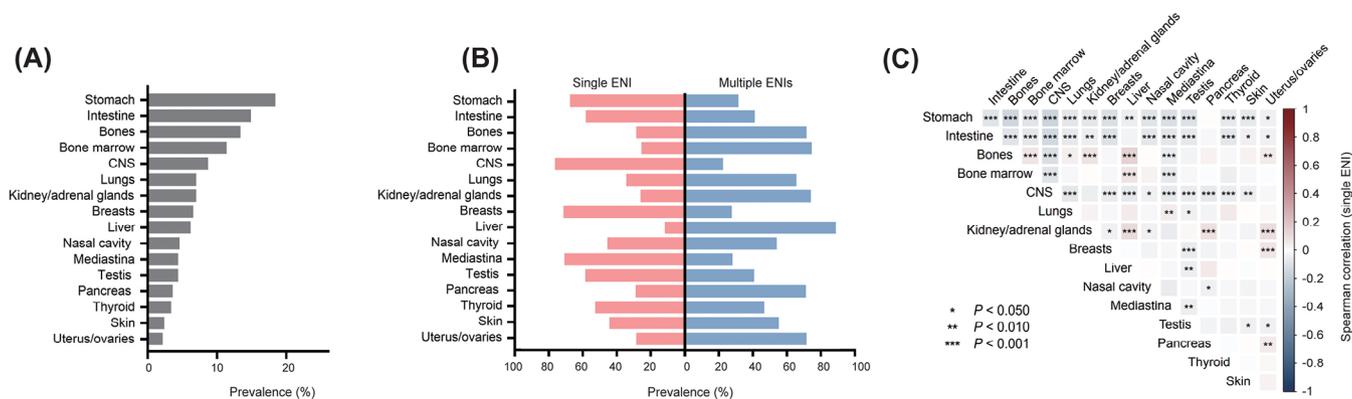


FIGURE 2 Distribution of involvement sites of the 4,784 patients with extranodal DLBCL. (A) Distribution of specific ENI sites in DLBCL patients. (B) Distribution of extranodal sites comparing single ENI and multiple ENIs. (C) Correlations between ENI sites. Red indicates a positive correlation between two sites of involvement, and blue indicates a negative correlation. Abbreviations: ENI, extranodal involvement; DLBCL, diffuse large B-cell lymphoma; CNS, central nervous system.

3.3 | Treatment and survival outcomes

The median follow-up period for the 4,784 patients was 40.5 months, ranging from 0.1 to 223.5 months. Patients with multiple ENIs exhibited significantly poorer outcomes compared to those with single ENI. Specifically, the 5-year PFS rates were 65.7% (95% CI, 62.6%-68.6%) for the single ENI group and 44.0% (95% CI, 39.2%-48.7%) for the multiple ENIs group; the 5-year OS rates were 74.3% (95% CI, 71.3%-77.1%) for the single ENI group and 55.3% (95% CI, 50.1%-60.6%) for the multiple ENIs group (Figure 3A). Prognosis of DLBCL patients of the single ENI group with different extranodal site involvement varied widely (Figure 3B). In the univariate analysis, involvements of the bones, bone marrow, kidney/adrenal glands, liver, CNS, and skin were associated with unfavorable prognosis in terms of both PFS and OS, whereas involvement of the stomach was linked to favorable outcomes. In the multivariate analysis, specific lymphoma involvement sites, including the bone marrow and CNS, remained as independent prognostic factors for both inferior OS and PFS after adjustment for standard prognostic variables (Table 2).

3.4 | Oncogenic mutations and genetic subtype-guided immunochemotherapy

We conducted an in-depth analysis of oncogenic mutations linked to DLBCL in 1,050 patients using gene panel sequencing. Notably, mutations in *MYD88*, CREB binding protein (*CREBBP*), interferon regulatory factor 4 (*IRF4*), and F-box and WD repeat domain containing 7 (*FBWX7*) were significantly more frequent in the multiple ENIs group than in the single ENI group, whereas *TET2* muta-

tions were more frequently observed in the single ENI group (Figure 4A). Gene mutation profiles exhibited significant variations across ENI sites. Patients with involvement of the GI tract exhibited an overall lower mutation rate than those with involvement of the breasts and testes, where mutation rates of both *Pim-1* proto-oncogene (*PIMI*) and *MYD88* exceeded 50% (Figure 4B-C). Among patients with GI tract involvement, mutations in *PIMI* (23.1% vs. 13.8%, $P < 0.001$) and *MYD88* (19.8% vs. 7.3%, $P = 0.002$) were more frequent in the multiple ENIs group than in the single ENI group, whereas mutations in *TET2* (11.2% vs. 20.8%, $P = 0.048$) were less frequent in the multiple ENIs group. Similarly, in patients with bone involvement, mutations in *CD79B* (12.3% vs. 27.6%, $P = 0.047$) and *TET2* (6.9% vs. 18.1%, $P = 0.050$) were less frequent in the multiple ENIs group. Among patients with testicular involvement, mutations in *KMT2D* (43.1% vs. 17.8%, $P = 0.017$) were more common in the multiple ENIs group, whereas in those with bone marrow invasion, mutations in *KMT2D* (14.2% vs. 37.8%, $P = 0.011$) were less frequent. Patients with lung involvement displayed lower *BTG1* mutations (2.4% vs. 38.2%, $P < 0.001$) in the multiple ENIs group than in the single ENI group.

A subset of patients ($n = 1,050$) with both the information on mutations of 35 genes and rearrangements of three genes (*BCL2*, *BCL6*, and *MYC*), was genetically classified using LymphPlex. The MCD-like subtype was notably more frequent in patients with multiple ENIs than in those with single ENI (17.8% vs. 10.2%, $P = 0.011$; Figure 4D). Significant correlations were observed among various mutations in extranodal DLBCL (Figure 4E). *PIMI* mutations exhibited strong co-occurrence with mutations in *MYD88*, BTG anti-proliferation factor 2 (*BTG2*), *CD79B*, macrophage-expressed 1 (*MPEG1*), and TBLIX/Y-related 1 (*TBLIXR1*), both in single and multiple ENIs patients.

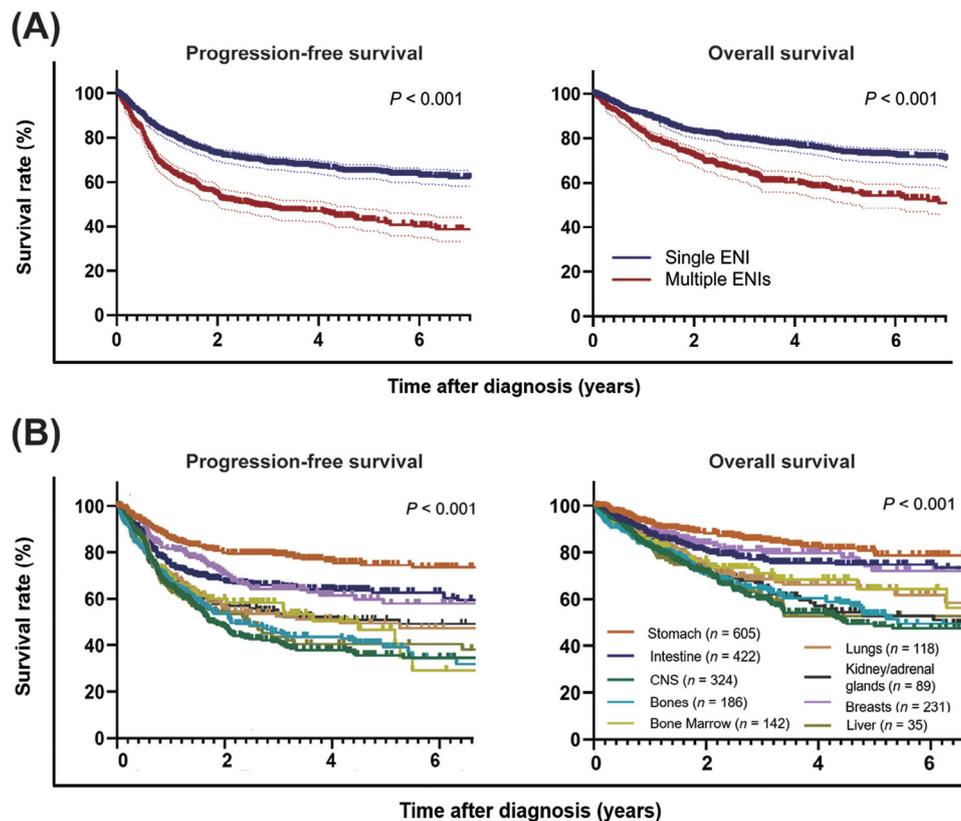


FIGURE 3 Prognosis of patients with extranodal DLBCL who received first-line treatment. (A) PFS and OS of patients with single ENI/multiple ENIs. (B) PFS and OS of the single ENI group according to sites of ENI. Abbreviations: DLBCL, diffuse large B-cell lymphoma; PFS, progression-free survival; OS, overall survival; ENI, extranodal involvement. CNS, central nervous system.

Additionally, *MYD88* mutations in single ENI group, 74.7% of which were the L265P variant (Supplementary Table S2), frequently co-occurred with mutations in *CD79B*, *BTG* anti-proliferation factor 1 (*BTG1*), *MPEGI1*, and *TBLIXR1*. Instances of mutual exclusivity between mutations were infrequent. Notably, in multiple ENIs group, *PIMI1* exhibited a significant negative correlation with *TP53*, while in single ENI group, *MYD88* displayed a significant negative correlation with TNF α induced protein 3 (*TNFAIP3*).

Based on the genetic subtypes of patients, we assessed the real-world efficacy of different treatment regimens on extranodal DLBCL. To overcome biases among variables, a PSM analysis was performed to assess treatment effects. The R-CHOP-X group demonstrated superior OS and PFS compared to the standard R-CHOP and R-CHOP-based group (Figure 4F). Specifically, patients receiving R-CHOP-X regimen achieved 5-year OS and PFS rates of 85.0% (95% CI, 80.6%-89.5%) and 72.1% (95% CI, 67.3%-76.7%), respectively, compared to 72.3% (95% CI, 67.0%-77.5%) for 5-year OS and 53.2% (95% CI, 48.5%-58.2%) for 5-year PFS in patients treated with standard R-CHOP and R-CHOP-based regimen.

3.5 | Tumor microenvironment alterations

To further elucidate the involvement of immune cells in ENI, we examined the activity scores of tumor immune cell recruitment using tumor immunophenotyping with RNA sequencing data from 434 patients. This cohort comprised 214 patients with multiple ENIs and 220 patients with single ENI. However, our analysis did not reveal any significant differences in activity scores among immune cell subsets between the two groups (Figure 5A).

GO and KEGG analyses revealed that several immune-associated signaling pathways, such as cytokine-mediated signaling pathways, cellular response to cytokine stimulus, response to cytokine, and immune response, were downregulated in patients with multiple ENIs, compared to those with single ENI (Figure 5B). This downregulation was further supported by our findings in the GSEA (Figure 5C). GO analysis indicated that T-cell immunity-related processes (T-cell differentiation, lymphocyte differentiation, and lymphocyte migration) and chemokine signaling (chemokine-mediated signaling, cellular response to chemokine, and cell chemotaxis) were

TABLE 2 Univariate and multivariate analyses for PFS and OS of the R-CHOP cohort ($n = 3,027$).

Variable	Univariate analysis				Multivariate analysis			
	PFS		OS		PFS		OS	
	P value	HR (95%CI)	P value	HR (95%CI)	P value	HR (95%CI)	P value	HR (95%CI)
ECOG score, ≥ 2	<0.001	2.31 (1.95-2.73)	<0.001	2.93 (2.42-3.54)	<0.001	1.52 (1.27-1.81)	<0.001	1.94 (1.59-2.36)
Ann Arbor stage, III-IV	<0.001	2.69 (2.29-3.17)	<0.001	3.12 (2.54-3.83)	<0.001	1.66 (1.37-2.00)	<0.001	1.78 (1.41-2.25)
Age, >60 years	<0.001	1.32 (1.15-1.53)	<0.001	1.93 (1.62-2.31)	0.017	1.20 (1.03-1.39)	<0.001	1.73 (1.45-2.08)
Elevated serum LDH	<0.001	3.09 (2.62-3.64)	<0.001	3.48 (2.85-4.25)	<0.001	2.15 (1.79-2.57)	<0.001	2.33 (1.87-2.90)
Gastric involvement	<0.001	0.45 (0.35-0.57)	<0.001	0.53 (0.40-0.71)	<0.001	0.57 (0.44-0.73)	0.012	0.69 (0.52-0.92)
Intestinal involvement	0.517	0.93 (0.76-1.15)	0.185	0.84 (0.65-1.09)	/	/	/	/
Bone involvement	<0.001	1.91 (1.63-2.25)	<0.001	1.85 (1.52-2.24)	0.030	1.21 (1.02-1.44)	0.338	1.11 (0.90-1.36)
Bone marrow involvement	<0.001	1.69 (1.40-2.04)	<0.001	1.68 (1.34-2.10)	0.008	1.30 (1.07-1.58)	0.045	1.27 (1.01-1.60)
CNS involvement	<0.001	2.49 (1.86-3.33)	<0.001	2.76 (1.91-4.00)	<0.001	2.22 (1.64-2.99)	<0.001	1.94 (1.32-2.84)
Lung involvement	0.026	1.31 (1.03-1.67)	0.193	1.21 (0.91-1.63)	0.958	1.01 (0.79-1.29)	/	/
Renal/adrenal gland involvement	0.019	1.32 (1.05-1.67)	0.001	1.54 (1.19-2.00)	0.620	0.94 (0.74-1.20)	0.803	1.03 (0.79-1.35)
Breast involvement	0.317	0.87 (0.66-1.14)	0.052	0.71 (0.50-1.00)	/	/	/	/
Liver involvement	<0.001	1.64 (1.28-2.09)	0.001	1.62 (1.21-2.17)	0.494	1.09 (0.85-1.40)	0.824	1.03 (0.77-1.39)
Nasal cavity involvement	0.542	1.12 (0.78-1.60)	0.512	0.84 (0.50-1.41)	/	/	/	/
Mediastinal involvement	0.003	1.54 (1.15-2.05)	0.804	1.05 (0.69-1.60)	0.031	1.40 (1.03-1.89)	/	/
Testicular involvement	0.413	0.87 (0.62-1.22)	0.685	0.92 (0.62-1.37)	/	/	/	/
Pancreatic involvement	0.521	1.10 (0.81-1.50)	0.029	1.43 (1.04-1.98)	/	/	0.467	1.13 (0.81-1.58)
Thyroid involvement	0.068	0.67 (0.43-1.03)	0.176	0.70 (0.42-1.17)	/	/	/	/
Skin involvement	0.003	1.69 (1.20-2.39)	0.008	1.75 (1.16-2.63)	0.150	1.30 (0.91-1.84)	0.153	1.36 (0.89-2.06)
Uterus/ovarian involvement	0.517	1.17 (0.73-1.86)	0.530	0.81 (0.42-1.56)	/	/	/	/

Abbreviations: PFS, progression-free survival; OS, overall survival; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone; HR, Hazard Ratio; CI, Confidence Interval; ECOG, Eastern Cooperative Oncology Group; LDH, lactate dehydrogenase; CNS, central nervous system.

downregulated in patients with multiple ENIs, compared to those with single ENI involving the GI tract, lungs, and breasts (Figure 5D).

The LME categories were applied to all patients with RNA sequencing data. Interestingly, the proportion of LME-MS was significantly lower in patients with multiple ENIs than those with single ENI. No significant differences were observed in the distribution of the other three LME categories. LME analysis for specific ENI suggested enrichment of LME-MS in DLBCL with GI tract involvement and LME-IN in DLBCL with testicular involvement (Figure 5E).

4 | DISCUSSION

We performed this large-scale retrospective analysis in the rituximab era from Asia, confirmed the poor prognosis of DLBCL with multiple ENIs and highlighted the molecular characteristics of extranodal DLBCL. Our results confirmed that patients with multiple ENIs exhibited a high frequency of *MYD88*, *TET2*, *CREBBP* mutations,

and increased MCD-like subtypes. Genetic subtype-guided immunochemotherapy showed favorable efficacy in subgroup analyses after propensity score matching with 5-year OS and PFS, providing practical guidance for clinical decision-making.

The distribution of ENI sites in our cohort was generally consistent with previous reports [14, 19, 28]. The GI tract was the most frequently involved site, followed by bones and bone marrow. Notably, we identified a positive correlation between bones and GI involvement and a negative correlation between GI and pancreatic involvement—an association not previously reported. Furthermore, our findings revealed that R-CHOP is adopted as the standard first-line treatment for extranodal DLBCL [12]. Methotrexate and dose-adjusted R-DA-EPOCH are employed as standard first-line treatment for PCNSL and PMBCL, respectively [12, 29]. Remarkably, our study further validated the potential benefits of applying genetic subtype-guided immunochemotherapy in routine clinical practice.

Based on data from the US registry, extranodal DLBCL originating from the GI tract, lungs, and liver/pancreas

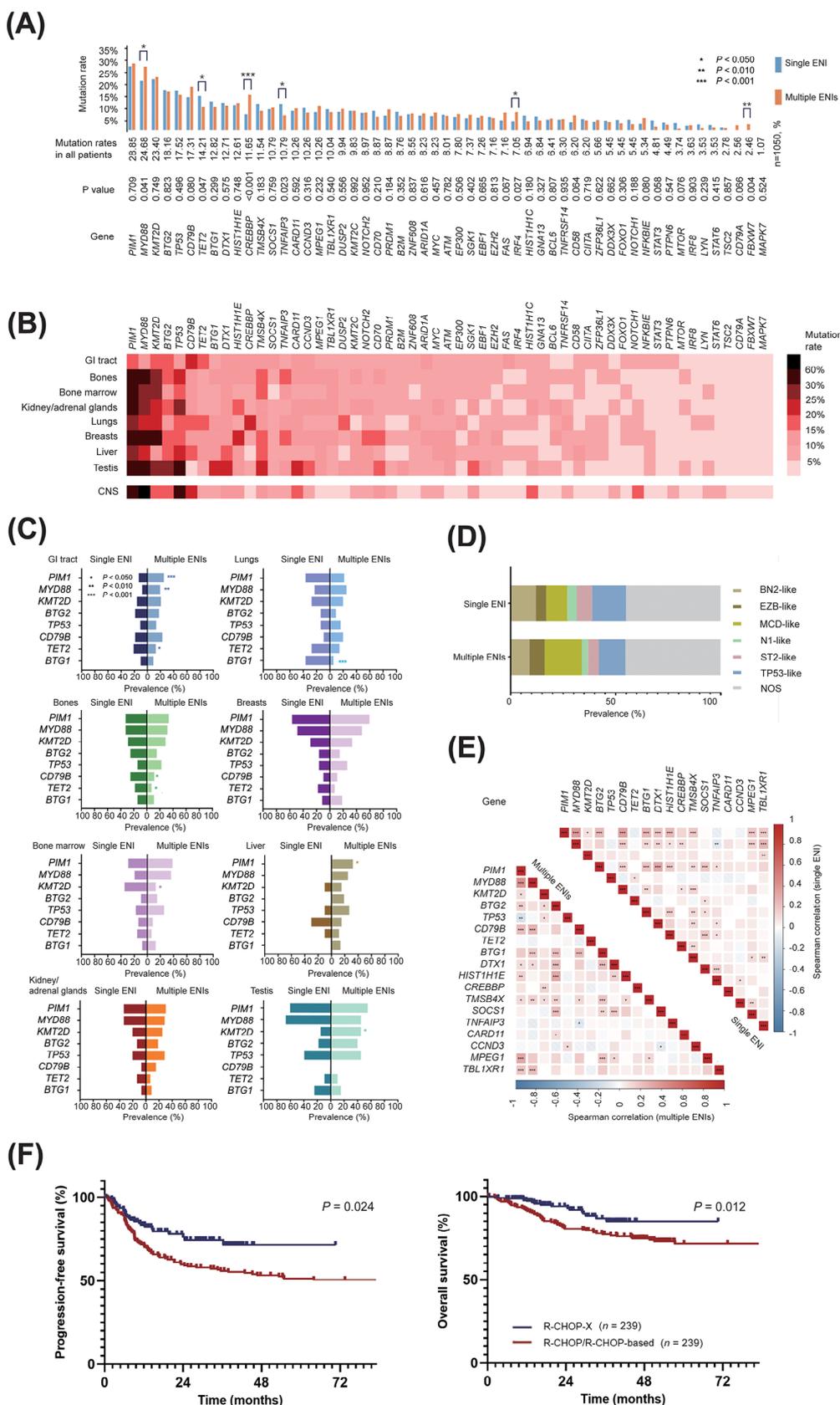


FIGURE 4 Relationship between oncogenic mutations and extranodal involvement in DLBCL. (A) Prevalence of genetic mutations in DLBCL patients with multiple ENI ($n = 372$) compared to those with single ENI ($n = 678$). (B) Mutation rates in the eight most common ENI sites and CNS. (C) Mutation rates in cases with involvement of the above sites were analyzed separately for patients with single ENI and multiple ENIs. The eight genes with the highest mutation frequencies are listed. (D) Prevalence of subtypes classified by LymphPlex in

indicate poor prognosis, while those arising in the head and neck region were associated with favorable outcomes [6]. Another extensive clinical analysis, involving 30,290 extranodal DLBCL patients, linked CNS, respiratory system, pancreas, and hepatobiliary involvements with poor prognosis [30]. A study in China, comprising 1,085 patients, reported that primary extranodal DLBCL from the stomach, breasts, sinus, lungs, and salivary glands have favorable outcomes, whereas those originating from CNS, testes, oral cavity, and kidney show poor prognosis [7]. In our study, GI tract, the most common ENI, was associated with favorable outcomes. In contrast, bone involvement emerged as an adverse prognostic factor, along with the bone marrow, CNS, liver, and lungs, as defined by NCCN-IPI. This observation was consistent with findings from another large cohort of 1221 DLBCL patients, indicating that the bones, spleen, kidney, and adrenal glands involvement, whether primary or secondary ENI, were associated with advanced stages and poor prognosis [31]. These variations could stem from regional differences, or disparities in the biological characteristics revealed by our large cohort study. More cases with complete clinical and molecular data should be accumulated in the future to yield more definitive results.

Moreover, genetic mutation patterns appeared to influence the survival outcomes of patients with extranodal DLBCL. Our data not only demonstrated a correlation between *MYD88* mutations and multiple ENIs [19], but also revealed the co-occurrence of several gene mutations, including *PIMI1*, *MYD88*, *CD79B*, *TBL1XR1*, *BTG1*, and *MPEGI*. These mutations are highly prevalent in the MCD subtype, and contribute to NF- κ B activation in a B cell receptor (BCR)-dependent manner [23]. Additionally, genetic characteristics varied in a site-specific manner. Lymphomas with single ENI in immune-privilege sites, such as the testes and CNS, exhibited a high frequency of *MYD88* mutations, consistent with previous studies on primary testicular lymphoma [32–34].

Promisingly, these genetic characteristics show promise as potential targets for therapeutic interventions. Our findings underscored that, even when genetic subtyping is not universally available for every patient, the enrichment of genetic signatures at specific lymphoma involvement sites can robustly predict the suitability of targeted agent

interventions. Notably, recent clinical trials, such as the PHOENIX trial, demonstrated the survival benefits of therapies like ibrutinib, targeting the BCR signaling pathway, in IPI 2-5, non-GCB patients under 60 years old [35]. Additionally, the POLARIX trial has shown that polatuzumab vedotin, targeting the surface antigen CD79B, when combined with rituximab, cyclophosphamide, doxorubicin, and prednisone (Pola-CHP), extended the survival of patients with IPI 2-5 [36]. These results indicated that novel targeted agents have the potential to improve the prognosis of extranodal DLBCL with specific genetic alterations. Given the molecular heterogeneity of extranodal DLBCL, our study retrospectively categorized cases for genetic subtyping and identified candidates eligible for the addition of targeted agents based on the GUIDANCE-01 study [26]. We explored the real-world outcomes of the R-CHOP-X group in a multicenter cohort and obtained encouraging results, suggesting that extranodal DLBCL patients receiving R-CHOP-X may achieve favorable outcomes. The EXPECT study thus presented a unique model for analyzing novel targeted therapies and offered valuable guidance for precise clinical treatment.

Tumor microenvironment plays a critical role in DLBCL progression [37]. Focusing on differential genes, downregulation in DLBCL with multiple ENIs related to granulocyte and lymphocyte chemotaxis and migration, T-cell activity, and cytokine-mediated immunity. These alterations may contribute to immune evasion and tumor dissemination in patients with multiple ENIs. Further analysis and validations using larger sequencing datasets are necessary to better understand how the tumor microenvironment influences extranodal invasion in DLBCL.

This study had some limitations. It primarily focused on the first-line treatment of DLBCL and did not provide systematic statistics on additional treatment modalities, such as radiotherapy, sequential autologous hematopoietic stem cell transplantation, and chimeric antigen receptor T-cell therapy. Additionally, this study was designed retrospectively, which may have introduced bias. Nevertheless, the strength of this study lied in its real-world, multicenter design, which minimized selection bias and provided extensive molecular insights into the survival outcomes of extranodal DLBCL in China.

patients with multiple ENIs and those with single ENI. (E) Correlations between common mutations (overall mutation rate >10%) and co-occurring mutations visualized in a correlation matrix. Bivariate correlation was calculated using Spearman's rank correlation. Asterisks indicate significant bivariate correlations ($P < 0.05$). Red indicates a positive correlation, and blue indicates a negative correlation. (F) PFS and OS of patients in the R-CHOP-X group and the R-CHOP/R-CHOP-based group. Abbreviations: GI, gastrointestinal; DLBCL, diffuse large B-cell lymphoma; ENI, extranodal involvement; CNS, central nervous system; R-CHOP, rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone.

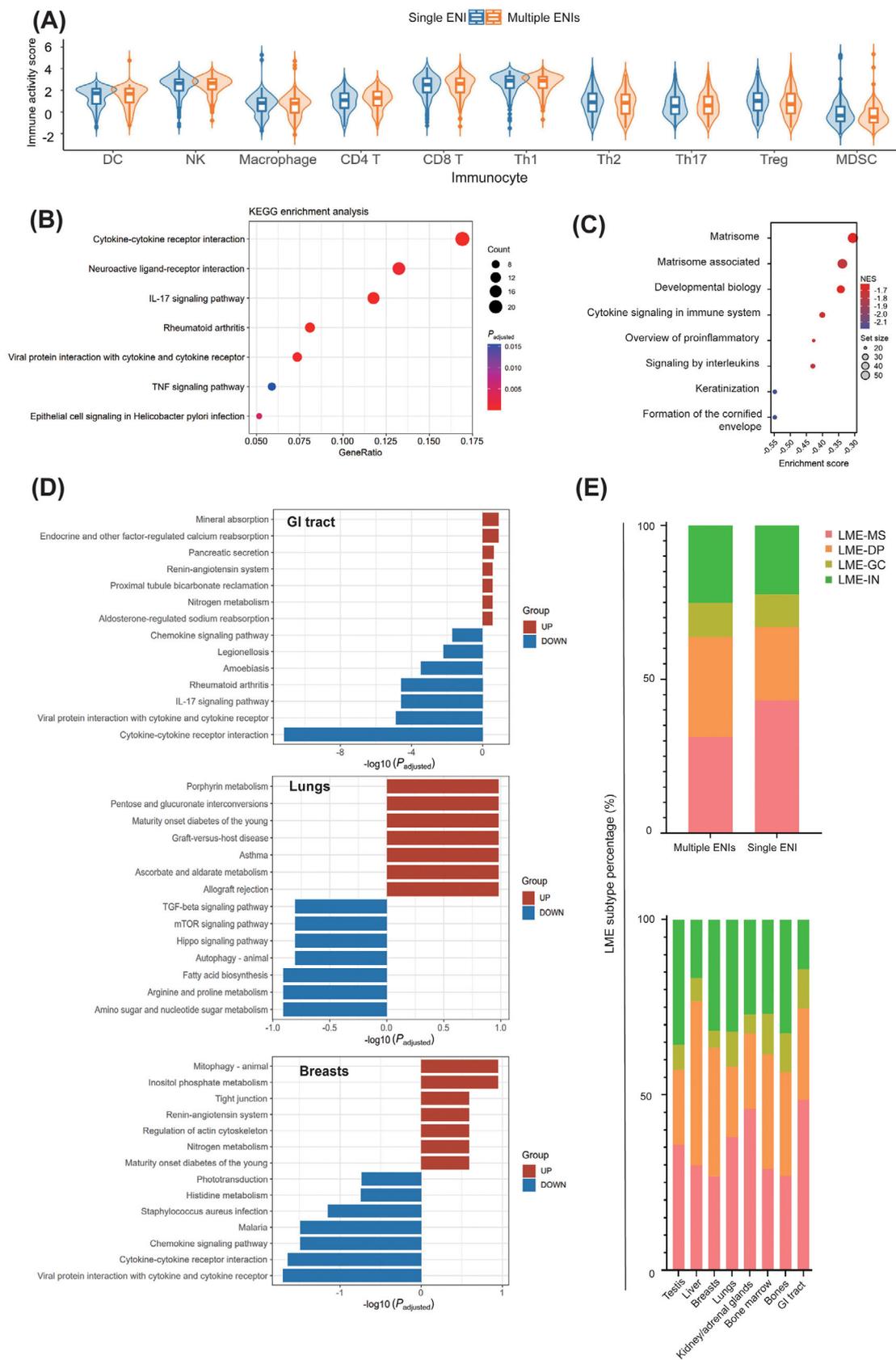


FIGURE 5 Relationship between intratumoral immune cells and ENI sites in DLBCL. (A) Immunity activity scores of indicated immune cells in patients with multiple ENIs ($n = 220$) and those with single ENI ($n = 214$). (B) KEGG terms in the multiple ENI group compared to the single ENI group. The color of points indicates the $-\log(\text{adjusted } P \text{ value})$ of dysregulated pathways in the two groups, and the size of points indicates the number of genes included in each gene set. (C) Downregulated GSEA terms in the multiple ENI group compared to the single

ENI group. The color of points indicates the $-\log(P \text{ value})$ of dysregulated pathways in the two groups, and the size of points indicates the number of genes included in each gene set. (D) Pathways downregulated at different extranodal sites in cases with single ENI and multiple ENIs. (E) Patients were categorized into four groups based on LME clusters, and the percentage of each group in patients with different numbers and locations of ENI was observed separately. Abbreviations: DLBCL, diffuse large B-cell lymphoma; ENI, extranodal involvement; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GSEA, Gene Set Enrichment Analysis; GI, gastrointestinal tract.

5 | CONCLUSIONS

This study shed light on the clinical and molecular characteristics of extranodal DLBCL in a multicenter real-world cohort. Multiple ENIs were associated with poor prognosis, and patients with multiple ENIs exhibited a higher frequency of *MYD88* and *CREBBP* mutations and an increased proportion of the MCD-like subtype. The favorable efficacy of the R-CHOP-X in extranodal DLBCL was preliminarily validated in this real-world cohort, and further basic research studies and prospective clinical trials are warranted to unravel the molecular mechanisms and clinical attributes of extranodal lymphoma, thereby guiding the selection of optimal clinical treatment strategies.

AFFILIATIONS

¹Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics, National Research Center for Translational Medicine at Shanghai, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, P. R. China

²Department of Hematology, Nanfang Hospital of Southern Medical University, Guangzhou, Guangdong, P. R. China

³Institute of Hematology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, P. R. China

⁴Clinical Research Institute, Shanghai Jiao Tong University School of Medicine, Shanghai, P. R. China

⁵Department of Lymphoma, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin, P. R. China

⁶Department of Hematology, the First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, P. R. China

⁷Department of Hematology, The First Affiliated Hospital of University of Science and Technology of China, Hefei, Anhui, P. R. China

⁸Department of Hematology, The Affiliated Cancer Hospital of Zhengzhou University, Henan Provincial Institute of Hematology, Zhengzhou, Henan, P. R. China

⁹Department of Hematology, Henan Provincial People's Hospital, Zhengzhou University, Zhengzhou, Henan, P. R. China

¹⁰National Clinical Research Center for Hematologic Disease, Peking University Institute of Hematology, Peking University Peoples Hospital, Beijing, P. R. China

¹¹Department of Hematology, Shandong Provincial Hospital, Shandong University, Jinan, Shandong, P. R. China

¹²Department of Hematology, Fujian Medical University Union Hospital, Fujian Institute of Hematology, Fujian Provincial Key Laboratory on Hematology, Fuzhou, Fujian, P. R. China

¹³Department of Hematology, Fujian Institute of Hematology, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian, P. R. China

¹⁴Department of Hematology, The First Affiliated Hospital of Fujian Medical University, Fuzhou, Fujian, P. R. China

¹⁵Department of Hematology, The First Hospital of Jilin University, Changchun, Jilin, P. R. China

¹⁶Department of Hematology, Nanjing Drum Tower Hospital, Nanjing, Jiangsu, P. R. China

¹⁷Department of Hematology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, P. R. China

¹⁸Department of Hematology, The Affiliated Hospital of Xuzhou Medical University, Xuzhou, Jiangsu, P. R. China

¹⁹Department of Hematology, The First People's Hospital of Yunnan Province, Kunming, Yunnan, P. R. China

²⁰Department of Hematology, General Hospital of Northern Theater Command, Shenyang, Liaoning, P. R. China

²¹Center of Hematology, The First Affiliated Hospital of Nanchang University, Nanchang, Jiangxi, P. R. China

²²Department of Hematology, Tongji Hospital, School of Medicine, Tongji University, Shanghai, P. R. China

²³Department of Lymphoma & Hematology, Hunan Cancer Hospital, The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, Hunan, P. R. China

²⁴Department of Hematology, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, P. R. China

²⁵Department of Hematology, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, P. R. China

²⁶Department of Hematology, The First Affiliated Hospital of Xiamen University, Xiamen, Fujian, P. R. China

²⁷Department of Hematology, Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, Zhejiang, P. R. China

²⁸Jiangsu Institute of Hematology, The First Affiliated Hospital of Soochow University, Suzhou, Jiangsu, P. R. China

²⁹Department of Medical Oncology, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong, P. R. China

³⁰Department of Hematology, Lanzhou University Second Hospital, Lanzhou, Gansu, P. R. China

³¹Department of Hematology, The Second Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, P. R. China

³²Department of Oncology, Affiliated Hospital of Nantong University, Nantong, Jiangsu, P. R. China

³³Department of Hematology, The First Affiliated Hospital, Harbin Medical University, Harbin, Heilongjiang, P. R. China

³⁴Department of Hematology, Hainan Cancer Hospital, Haikou, Hainan, P. R. China

³⁵Department of Hematology, Xinqiao Hospital, Third Military Medical University, Chongqing, P. R. China

³⁶Department of Hematology, General Hospital of Ningxia Medical University, Yinchuan, Ningxia, P. R. China

³⁷Department of Hematology, Beijing Hospital, Ministry of Health, Beijing, P. R. China

³⁸Department of Hematology, The Fifth Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong, P. R. China

³⁹Department of Hematology, Anshan Central Hospital Affiliated to China Medical University, Anshan, Liaoning, P. R. China

⁴⁰Department of Hematology and Oncology, Harbin Institute of Hematology and Oncology, Harbin, Heilongjiang, P. R. China

⁴¹Department of Hematology, West China Hospital, Sichuan University, Chengdu, Sichuan, P. R. China

⁴²Pôle de Recherches Sino-Français en Science du Vivant et Génomique, Laboratory of Molecular Pathology, Shanghai, P. R. China

AUTHOR CONTRIBUTIONS

Si-Yuan Chen and Peng-Peng Xu collected and analyzed the data, and contributed to writing the article. Ru Feng, Guo-Hui Cui, Peng-Peng Xu, Li Wang, Shu Cheng, Hui-Lai Zhang, Xiao-Lei Wei, Yong-Ping Song, Kai-Yang Ding, Li-Hua Dong, Zun-Min Zhu, Shen-Miao Yang, Xin Wang, Ting-Bo Liu, Jian-Da Hu, Xiao-Yun Zheng, Ou Bai, Jing-Yan Xu, Liang Huang, Wei Sang, Ke-Qian Shi, Fan Zhou, Fei Li, Ai-Bin Liang, Hui Zhou, Si-Guo Hao, Hong-Hui Huang, Bin Xu, Wen-Bin Qian, Cai-Xia Li, Zhi-Ming Li, Chong-Yang Wu, Xiao-Bo Wang, Wen-Yu Shi, Shu-Ye Wang, Yu-Yang Tian, Xi Zhang, Ke-Shu Zhou, Li-Juan Cui, Hui Liu, Huo Tan, Qing Leng, Dong-Lu Zhao, Ting Niu recruited patients, gathered detailed clinical information and provided data management for the study. Ru Feng and Guo-Hui Cui provided executive active data surveillance and performed the patient selection process. Si-Yuan Chen was responsible for bioinformatics investigation. Peng-Peng Xu and Rong-Ji Mu provided technical support. Wei-Li Zhao and conceptualized the study, provided guidance and oversight for the research, and contributed to manuscript writing. All authors reviewed the manuscript critically and approved the content.

ACKNOWLEDGMENTS

We express our gratitude to the dedicated physicians who enrolled patients and to all the patients who consented to the analysis of their clinical data.

CONFLICT OF INTEREST STATEMENT

We declare no competing interests.

FUNDING INFORMATION

This study received support from Shanghai Clinical Research Center for Cell Therapy (23J41900100), National Natural Science Foundation of China (82130004), Clinical Research Plan of Shanghai Hospital Development Center (SHDC2020CR1032B), National Key Research and Development Program (2022YFC2502600), and Multicenter Clinical Research Project by Shanghai Jiao Tong University School of Medicine (DLY201601).

The sponsors had no role in study design, data collection, analysis, interpretation, or writing of the report.

DATA AVAILABILITY STATEMENT

Genomic and gene expression data have been deposited on <https://www.biosino.org> (OEP001143). All data have been deposited in a data repository specific to this study (<https://edc.trialdata.cn>) and are accessible to researchers meeting the criteria for data access. For proposals requesting individual participant data that underlie the results reported in this article (after de-identification), a steering committee involving all principal investigators will evaluate the request and make the decision before sending the database to academic partners. For data requests, please contact zhao.weili@yahoo.com. There are currently no plans to share data not included in this paper.

The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

ETHICAL APPROVAL AND CONSENT TO PARTICIPATE

This study was approved by the Ethics Committee of Shanghai Ruijin Hospital (2022-91), with informed consent obtained from all patients in accordance with the Declaration of Helsinki. This trial was registered at ClinicalTrials.gov, number NCT05907447.

ORCID

Rong-Ji Mu  <https://orcid.org/0000-0002-7525-7930>

Xin Wang  <https://orcid.org/0000-0001-8051-1481>

Jian-Da Hu  <https://orcid.org/0000-0002-4438-2544>

Wei Sang  <https://orcid.org/0000-0002-1170-7074>

Fei Li  <https://orcid.org/0000-0002-7765-5975>

Bin Xu  <https://orcid.org/0000-0002-7271-4438>

Wei-Li Zhao  <https://orcid.org/0000-0002-6834-1616>

REFERENCES

1. A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project. *Blood*. 1997;89(11):3909–18.

2. Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375–90.
3. Kridel R, Telio D, Villa D, Sehn LH, Gerrie AS, Shenkier T, et al. Diffuse large B-cell lymphoma with testicular involvement: outcome and risk of CNS relapse in the rituximab era. *Br J Haematol*. 2017;176(2):210–21.
4. Roschewski M, Phelan JD, Jaffe ES. Primary Large B-cell Lymphomas of Immune-Privileged Sites. *Blood*. 2024;144(25):2593–603.
5. Bobillo S, Joffe E, Lavery JA, Sermer D, Ghione P, Noy A, et al. Clinical characteristics and outcomes of extranodal stage I diffuse large B-cell lymphoma in the rituximab era. *Blood*. 2021;137(1):39–48.
6. Castillo JJ, Winer ES, Olszewski AJ. Sites of extranodal involvement are prognostic in patients with diffuse large B-cell lymphoma in the rituximab era: an analysis of the Surveillance, Epidemiology and End Results database. *Am J Hematol*. 2014;89(3):310–4.
7. Shi Y, Han Y, Yang J, Liu P, He X, Zhang C, et al. Clinical features and outcomes of diffuse large B-cell lymphoma based on nodal or extranodal primary sites of origin: Analysis of 1,085 WHO classified cases in a single institution in China. *Chin J Cancer Res*. 2019;31(1):152–61.
8. Savage KJ. Primary mediastinal large B-cell lymphoma. *Blood*. 2022;140(9):955–70.
9. 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(1):29.
10. Hristov AC, Tejasvi T, Wilcox RA. Cutaneous B-cell lymphomas: 2023 update on diagnosis, risk-stratification, and management. *Am J Hematol*. 2023;98(8):1326–32.
11. Fu K, Weisenburger DD, Choi WWL, Perry KD, Smith LM, Shi X, et al. Addition of rituximab to standard chemotherapy improves the survival of both the germinal center B-cell-like and non-germinal center B-cell-like subtypes of diffuse large B-cell lymphoma. *J Clin Oncol*. 2008;26(28):4587–94.
12. Sehn LH, Salles G. Diffuse Large B-Cell Lymphoma. *N Engl J Med*. 2021;384(9):842–58.
13. Ferry JA. Extranodal lymphoma. *Arch Pathol Lab Med*. 2008;132(4):565–78.
14. Ollila TA, Olszewski AJ. Extranodal Diffuse Large B Cell Lymphoma: Molecular Features, Prognosis, and Risk of Central Nervous System Recurrence. *Curr Treat Options Oncol*. 2018;19(8):38.
15. Kirkegaard MM, Rasmussen PK, Coupland SE, Esmali B, Finger PT, Graue GF, et al. Conjunctival Lymphoma—An International Multicenter Retrospective Study. *JAMA Ophthalmol*. 2016;134(4):406–14.
16. Yin X, Xu A, Fan F, Huang Z, Cheng Q, Zhang L, et al. Incidence and Mortality Trends and Risk Prediction Nomogram for Extranodal Diffuse Large B-Cell Lymphoma: An Analysis of the Surveillance, Epidemiology, and End Results Database. *Front Oncol*. 2019;9:1198.
17. A predictive model for aggressive non-Hodgkin's lymphoma. *New Engl J Med*. 1993;329(14):987–94.
18. Zhou Z, Sehn LH, Rademaker AW, Gordon LI, Lacasce AS, Crosby-Thompson A, et al. An enhanced International Prognostic Index (NCCN-IPI) for patients with diffuse large B-cell lymphoma treated in the rituximab era. *Blood*. 2014;123(6):837–42.
19. Shen R, Xu P-P, Wang N, Yi H-M, Dong L, Fu D, et al. Influence of oncogenic mutations and tumor microenvironment alterations on extranodal invasion in diffuse large B-cell lymphoma. *Clin Transl Med*. 2020;10(7):e221.
20. Kim YR, Kim JS, Min YH, Hyunyon D, Shin H-J, Mun Y-C, et al. Prognostic factors in primary diffuse large B-cell lymphoma of adrenal gland treated with rituximab-CHOP chemotherapy from the Consortium for Improving Survival of Lymphoma (CISL). *J Hematol Oncol*. 2012;5:49.
21. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandembroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007;370(9596):1453–7.
22. Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol*. 2014;32(27):3059–68.
23. Schmitz R, Wright GW, Huang DW, Johnson CA, Phelan JD, Wang JQ, et al. Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. *New Engl J Med*. 2018;378(15):1396–407.
24. Shen R, Fu D, Dong L, Zhang M-C, Shi Q, Shi Z-Y, et al. Simplified algorithm for genetic subtyping in diffuse large B-cell lymphoma. *Signal Transduct Target Ther*. 2023;8(1):145.
25. Wright GW, Huang DW, Phelan JD, Coulibaly ZA, Roulland S, Young RM, et al. A Probabilistic Classification Tool for Genetic Subtypes of Diffuse Large B Cell Lymphoma with Therapeutic Implications. *Cancer Cell*. 2020;37(4):551–68.
26. Zhang M-C, Tian S, Fu D, Wang L, Cheng S, Yi H-M, et al. Genetic subtype-guided immunochemotherapy in diffuse large B cell lymphoma: The randomized GUIDANCE-01 trial. *Cancer Cell*. 2023;41(10):1705–16.
27. Kotlov N, Bagaev A, Revuelta MV, Phillip JM, Cacciapuoti MT, Antysheva Z, et al. Clinical and Biological Subtypes of B-cell Lymphoma Revealed by Microenvironmental Signatures. *Cancer Discov*. 2021;11(6):1468–89.
28. Li X, Liu Z, Cao J, Hong X, Wang J, Chen F, et al. Rituximab in combination with CHOP chemotherapy for the treatment of diffuse large B cell lymphoma in China: a 10-year retrospective follow-up analysis of 437 cases from Shanghai Lymphoma Research Group. *Ann Hematol*. 2012;91(6):837–45.
29. Calimeri T, Steffanoni S, Gagliardi F, Chiara A, Ferreri AJM. How we treat primary central nervous system lymphoma. *ESMO Open*. 2021;6(4):100213.
30. Gupta V, Singh V, Bajwa R, Meghal T, Sen S, Greenberg D, et al. Site-Specific Survival of Extra Nodal Diffuse Large B-Cell Lymphoma and Comparison With Gastrointestinal Diffuse Large B-Cell Lymphoma. *J Hematol*. 2022;11(2):45–54.
31. Takahashi H, Tomita N, Yokoyama M, Tsunoda S, Yano T, Murayama K, et al. Prognostic impact of extranodal involvement in diffuse large B-cell lymphoma in the rituximab era. *Cancer*. 2012;118(17):4166–72.
32. Deng L, Xu-Monette ZY, Loghavi S, Manyam GC, Xia Y, Visco C, et al. Primary testicular diffuse large B-cell lymphoma displays distinct clinical and biological features for treatment failure in

- rituximab era: a report from the International PTL Consortium. *Leukemia*. 2016;30(2):361–72.
33. Oishi N, Kondo T, Nakazawa T, Mochizuki K, Tanioka F, Oyama T, et al. High prevalence of the MYD88 mutation in testicular lymphoma: Immunohistochemical and genetic analyses. *Pathol Int*. 2015;65(10):528–35.
34. Kraan W, van Keimpema M, Horlings HM, Schilder-Tol EJM, Oud MECM, Noorduyn LA, et al. High prevalence of oncogenic MYD88 and CD79B mutations in primary testicular diffuse large B-cell lymphoma. *Leukemia*. 2014;28(3):719–20.
35. Younes A, Sehn LH, Johnson P, Zinzani PL, Hong X, Zhu J, et al. Randomized Phase III Trial of Ibrutinib and Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone in Non-Germinal Center B-Cell Diffuse Large B-Cell Lymphoma. *J Clin Oncol*. 2019;37(15):1285–95.
36. Tilly H, Morschhauser F, Sehn LH, Friedberg JW, Trněný M, Sharman JP, et al. Polatuzumab Vedotin in Previously Untreated Diffuse Large B-Cell Lymphoma. *New Engl J Med*. 2022;386(4):351–63.
37. Scott DW, Gascoyne RD. The tumour microenvironment in B cell lymphomas. *Nat Rev Cancer*. 2014;14(8):517–34.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Chen S-Y, Xu P-P, Feng R, Cui G-H, Wang L, Cheng S, et al. Extranodal diffuse large B-cell lymphoma: Clinical and molecular insights with survival outcomes from the multicenter EXPECT study. *Cancer Commun.* 2025;1–17. <https://doi.org/10.1002/cac2.70033>