Lymphoma Prognosis with Lesion-Anatomy Context Fusion and Attention-Based Multi-Lesion Aggregation

Song Zhang[1], Jiajin Zhang[1,2,3] *, Liheng Qiu[4], Wei Liu[1], Dakai Jin[1], Le Lu[1], Shenmiao Yang[4], and Ke Yan[1]

[1] DAMO Academy, Alibaba Group, Hangzhou, China
[2] Shanghai Jiao Tong University, Shanghai, China
[3] Hupan Lab, 310023, Hangzhou, China
[4] Peking University People's Hospital, Beijing, China

Abstract. Early identification of lymphoma patients with poor prognosis is crucial to determining personalized treatment plans and improving prognosis. Currently, commonly used prognostic biomarkers include clinical variables such as International Prognostic Index. Quantitative parameters based on PET/CT and deep learning methods have also shown promising results. However, there are still several challenges in PET/CT-based prognostic studies: heterogeneity in the number and location of lesions, insufficient representation of lesion features, and the lack of anatomical context modeling of the lesions. We propose a novel framework named **LAMP**, with lesion-anatomy context fusion and attention-based multi-lesion aggregation as its two key components. The former takes into account information about the surrounding anatomical organs of the lesions to improve their representation. The latter treats each lesion region as an instance, assigning attention scores that reflect the contribution of each lesion, and aggregates them accordingly. A total of 229 lymphoma patients were collected to evaluate our model. In prediction tasks for progression-free survival and overall survival, the 5-fold cross-validation C-index is 0.791 and 0.828, respectively, outperforming existing models based on clinical variables and deep learning. **LAMP** has the potential to become a clinical auxiliary tool to differentiate patients with varying risk levels, facilitating the development of personalized treatment plans.

Keywords: PET/CT · Lymphoma Prognosis · Anatomical Context

1 Introduction

Lymphoma is a malignant neoplasm of the lymphatic system [1,2]. Unlike solid tumors, lymphoma often spread throughout the entire body. Diffuse large B-cell lymphoma (DLBCL), the most common type, accounts for 30% of non-Hodgkin's

---

* Corresponding author: Jiajin Zhang (zhangjiajin.zjj@alibaba-inc.com).

lymphoma [3,4]. Approximately 60% to 80% of patients achieve complete remission following first-line treatment with R-CHOP [5,6]. However, 20% to 40% may experience relapse or progression after initial therapy. Early identification of patients with poor prognosis is crucial, as it enables clinicians to avoid potentially ineffective therapy and timely adjust in treatment strategy [7,8,9].

Currently, the primary prognostic marker for DLBCL is the International Prognostic Index (IPI) and its variants [10,11]. IPI has long been the cornerstone for risk stratification in newly diagnosed DLBCL patients. However, IPI is predominantly based on clinical factors, overlooking tumor heterogeneity among individuals. Furthermore, with the development of new therapies, IPI has become increasingly inadequate to meet current clinical needs. As an essential clinical imaging tool for DLBCL prognosis, ^{18}F-FDG PET/CT provides a non-invasive method to assess tumor metabolism and capture tumor heterogeneity from a macro perspective. Prognostic biomarkers derived from PET/CT, such as metabolic tumor volume, and total lesion glycolysis, have shown promising results [14,15,16,17]. However, these metabolic parameters primarily rely on simple features, such as standard uptake value (SUV) and lesion volume, without fully capturing the complexity of tumor characteristics.

In addition to PET/CT-based quantitative parameters specifically designed for lymphoma, many other methods have been explored for prognosis and diagnosis. A common approach is radiomics feature extraction from regions of interest (ROIs) via manual or semi-automatic segmentation [18,19,20,12,13]. However, delineating ROIs is time-consuming, and general-purpose radiomics features may fail to provide specific and sufficient feature representation of lymphoma lesion regions. Recently, the use of deep learning techniques for PET/CT image analysis has grown [21,22,23]. However, these approaches of cropping fixed organ regions limit its application in diseases like lymphoma, where lesions are scattered in various anatomical regions, with heterogeneous lesion numbers, locations, and sizes across patients. Some researchers have attempted to apply the aforementioned traditional methods to lymphoma [24,25,26]. Meanwhile, Liu et al. [27] employed multi-task learning by integrating segmentation and prognosis prediction within a unified framework. However, they resampled whole-body PET/CT to a fixed-size input, which cannot adapt to varying lesion locations and sizes. Graph-based approaches [28,29] provide a novel perspective for analyzing DLBCL. In these methods, each lesion is treated as a node of a graph, with radiomics features extracted as the attributes of each node. Graph attention mechanisms are then employed to model the interactions between lesions. However, these methods mainly focus on local lesion features while neglecting crucial spatial context, including lesion distribution and its anatomical correlation with surrounding organs.

For DLBCL, PET/CT-based prognostic assessment still presents several challenges: **Heterogeneity in lesion number and location**. Unlike solid tumors, where the ROI of a specific lesion/organ can be cropped for analysis, DLBCL lesions are dispersed across different regions and their size varies, making it difficult to define a standardized ROI. The inherent characteristic of lymphoma necessitates a more flexible approach to analysis. **Insufficient representation of lesion regions**. Radiomics features lack specificity in representing lesion regions, while whole-image-based deep learning methods usually focus on dominant lesions and may overlook small metastatic lesions. The strategy of extracting features from each lesion region separately and then aggregating them (usually through average pooling) also requires further optimization. **Lack of modeling the anatomical context of lesions**. For example, extranodal involvement in organs like the liver and marrow in DLBCL is typically associated with poorer prognosis [30], but current methods cannot explicitly model the anatomical context of lesions.

To address the aforementioned challenges, this paper proposes a model for **L**ymphoma using **A**natomical context to lesions and **M**ulti-lesion aggregation for **P**rognosis (**LAMP**). This approach aims to enhance the accuracy of prognosis by effectively extracting and integrating features from multiple lesions while considering the anatomy of the lesions. The contributions of this work are as follows:

- We use attention-based multi-lesion aggregation to solve the issue of heterogeneous lesion distribution. The learned attention score is interpretable and can show how each lesion location contributes to prognosis prediction.
- We characterize the interaction between lesions and anatomical structures to enhance the anatomical context information of each lesion, thereby improving the representation of the lesion in its spatial and functional environment, which contributes to more accurate prognosis prediction.
- Extensive validation demonstrates that our approach outperforms existing prognosis models. Both quantitative evaluations and qualitative analyses confirmed the advantages of the proposed method.

2 Method

The overview of the proposed framework is presented in Fig. 1. For PET/CT-based prognostic prediction in lymphoma, we implement a comprehensive computational pipeline comprising three stages. Initially, a multimodal segmentation network is employed to achieve voxel-level lesion delineation, meanwhile generating high-dimensional feature maps that maintain spatial correspondence with the original imaging data. Subsequently, whole-body anatomical segmentation is obtained on CT volumes to derive organ masks, enabling lesion-anatomy context fusion through an attention-based feature interaction mechanism. Finally, multi-lesion aggregation is implemented to integrate features across distributed lesion regions, incorporating both local lesion characteristics and contextual organ information for precise prognostic stratification. We denote the N training data as $\mathbb{D} = \{X^i_{\mathrm{CT}}, X^i_{\mathrm{PET}}, C^i, S^i\}_{i=1}^N$, where X^i_{CT}, and X^i_{PET} denote the input CT image and PET image, C^i and S^i denote censorship status and survival time. Now we will delve into further details.

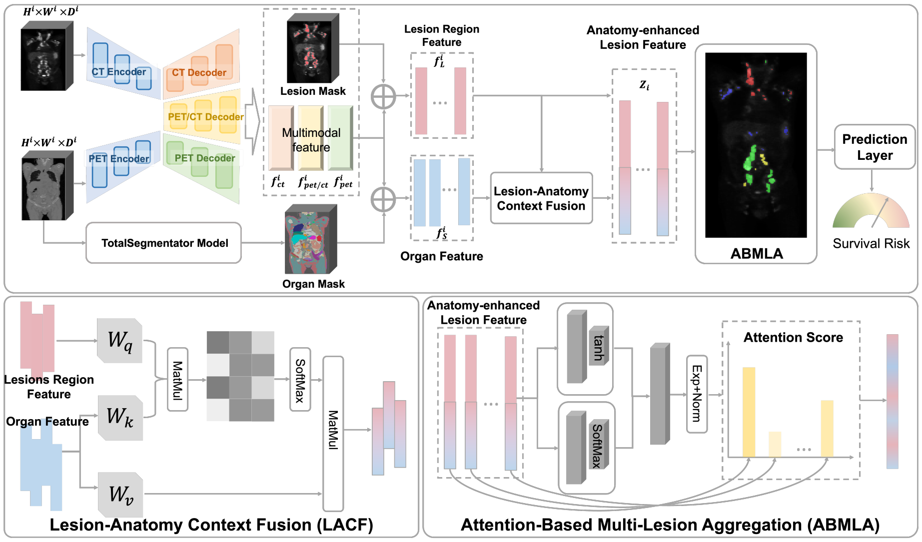

Fig. 1. Overview architecture of our proposed LAMP model.

2.1 Lesion Segmentation and Feature Extraction

Based on the nnUNet [31] framework, we designed a multimodal segmentation network that integrates PET and CT images. This network consists of two encoders with shared weights, each inputting either X_{CT} or X_{PET}. Following the encoders, three distinct decoders are introduced to process the feature representations from X_{CT}, X_{PET}, and the fusion of X_{CT} and X_{PET} features. Each decoder serves a distinct purpose: one extracts anatomical information from CT, another captures metabolic information from PET, and the third enables a deep interaction between the PET and CT information. Each decoder predicts the lesion masks separately. Finally, a pixel-wise weight map is computed to fuse the three probability maps adaptively and generate the final lesion mask. This design combines PET and CT in both feature and decision levels. On our DLBCL dataset of 229 cases under five-fold cross-validation, this lesion segmentation network achieved a Dice coefficient of 0.9199, a false positive volume (FPV) of 3.4940, and a false negative volume (FNV) of 2.0300. For lesion-level detection, it achieved a precision of 0.8931 and recall of 0.8991, significantly outperforming the plain nnUNet. It provides a solid foundation for our subsequent analysis.

To accurately differentiate lesions from normal tissues in PET/CT, a segmentation network needs to comprehensively encode each lesion's local appearance and global anatomical information. Therefore, the feature map of the segmentation network should also be valuable for downstream tasks, such as prognosis. Based on this assumption, we extracted feature maps with the same spatial dimensions as the input images from the final block of each decoder branch in the aforementioned segmentation network, and then pool lesion features from them. Formally, let $\mathbb{L}^i = \{l_1^i, l_2^i, \cdots, l_{i_J}^i\} \in \mathbb{R}^{H^i \times W^i \times D^i}$ represents the lesion regions predicted by the segmentation network for the i-th patient, and $\{f_{ct}^i \in \mathbb{R}^{32 \times H^i \times W^i \times D^i}, f_{pet}^i \in \mathbb{R}^{32 \times H^i \times W^i \times D^i}, f_{pet/ct}^i \in \mathbb{R}^{64 \times H^i \times W^i \times D^i}\}$ denote the multimodal feature set, where i_J is the overall number of lesions for the i-th patient and H^i,W^i,D^i are the height, width, and depth of X_{CT}^i and X_{PET}^i. The features of the j-th lesion on the i-th patient can be represented as:

$$f_{i_j}^i = Concat\left[f_{pet/ct}^i \oplus l_{i_j}^i, f_{pet}^i \oplus l_{i_j}^i, f_{ct}^i \oplus l_{i_j}^i\right] \tag{1}$$

where $\oplus$ means average pooling within the lesion mask.

2.2 Lesion-Anatomy Context Fusion

DLBCL may not only occur in lymph nodes, but also involve extranodal organs, with extranodal involvement often being associated with the poor prognosis of DLBCL. We aim to explore the relationship between prognosis and lesion involvement in extranodal organs by Lesion-Anatomy Context Fusion (LACF). Specifically, we characterize extranodal involvement by assessing the similarity between the features of the lesion regions and the features of the organ regions. This allows us to capture the degree of organ involvement of each lesion. Then, we weight and aggregate these organ region features to enrich the characterization and representation of each lesion region. This approach not only enhances the lesion feature's anatomical and metabolic profiles but also incorporates crucial information regarding the organ context, providing a more comprehensive understanding of the disease's progression and prognosis.

First, we use TotalSegmentator [32] to segment major anatomical structures $\mathbb{S}^i = \{s_1^i, s_2^i, \cdots, s_K^i\}$ on the CT images (The regions belonging to the same anatomical structure on the same side will be merged, reducing the number of labels from 117 to 71.). Following the same feature extraction method as described for the lesion regions, we extract the feature representation for each anatomical organ.

$$f_{s_k}^i = Concat\left[f_{pet/ct}^i \oplus s_k^i, f_{pet}^i \oplus s_k^i, f_{ct}^i \oplus s_k^i\right] \tag{2}$$

where $f_{s_k}^i$ stands for the features extracted from the s_k-th anatomical organs of the i-th patient. Next, we employ the attention mechanism to perform the association analysis between the lesion region features and the anatomical organs features. We use the features of the lesion region $f_L^i \in \mathbb{R}^{i_J \times 128}$ as query Q, the anatomical organs' feature representation $f_S^i \in \mathbb{R}^{K \times 128}$ as key K and value V:

$$\begin{aligned} h_L^i &= \mathrm{Attention}\left(Q, K, V\right) \\ &= \mathrm{Attention}\left(f_L^i W_q, f_S^i W_k, f_S^i W_v\right) \\ &= \mathrm{softmax}\left(\frac{(f_L^i W_q)(f_S^i W_k)^T}{\sqrt{d_k}}\right) f_S^i W_v. \end{aligned} \tag{3}$$

The attention mechanism calculates the similarity between the query and key, allowing the model to focus on the most relevant anatomical features that interact with the lesion. This enables the model to establish a more precise correlation between the lesion's characteristics and the surrounding anatomical context, thus improving the overall lesion representation and enhancing the prediction of prognosis. After that, $f^i_{i_j}$ and the i_j-th row from h^i_L are concatenated to get the anatomy-enhanced lesion feature

$$z^i_{i_j} = Concat\left[f^i_{i_j}, \left(h^i_L\right)_{i_j}\right]. \tag{4}$$

2.3 Attention-Based Multi-Lesion Aggregation

DLBCL presents unique challenges due to its heterogeneous anatomical distribution and variable number of lesions across patients. Conventional methodologies employed fixed-size whole-image input paradigms or naive average pooling of all lesion features. They are unable to adaptively process lesion regions with varying anatomical locations and pathological significance. To overcome these limitations, we propose a novel solution by reformulating the aggregation of the lesion feature as an attention-based multi-lesion aggregation (ABMLA) problem using gated attention mechanisms.

$$P^i = \sum_{j=1}^{i_J} a^i_{i_j} z^i_{i_j}, \quad a^i_{i_j} = \frac{exp\left\{w^T\left(tanh(V{z^i_{i_j}}^T) \oplus sigm(U{z^i_{i_j}}^T)\right)\right\}}{\sum\limits_{j=1}^{i_J} exp\left\{w^T\left(tanh(V{z^i_{i_j}}^T) \oplus sigm(U{z^i_{i_j}}^T)\right)\right\}} \tag{5}$$

where $a^i_{i_j}$ is the attention score learned by self-attention layers with parameters U,V,and w. Through that, the relative contribution of individual lesion regions to the predictive outcome could be assessed. P^i will be input into the linear layer to obtain the final survival risk score. NLL (negative log-likelihood) survival loss is leveraged to train the model [33].

3 Experiments

Data Collection. To evaluate the performance of our method, 229 pre-treatment whole-body PET/CT image pairs from patients with DLBCL are collected, with prognostic information available for 195 patients. This study focuses on two main outcomes: progression-free survival (PFS) and overall survival (OS). Compared to OS, PFS is more critical, as it is of primary concern to clinicians due to its direct relevance to disease progression in patients. Multiple preprocessing steps are applied to the PET/CT image pairs for model training to ensure proper alignment and standardization, including standardized uptake values conversion, registering the CT images to the PET images, and z-score normalization.

Implementation and Evaluation. Since the major focus of this paper is prognosis instead of lesion segmentation, we will mainly introduce the prognosis algorithm in this section. We used the Adam optimizer with an initial learning rate of 0.0001 to train our model. We utilized a batch size of 1 with a gradient accumulation step of 32 due to nonuniform lesion numbers across patients.

Table 1. Comparison results using C-index and Time-AUC metrics under five-fold cross-validation.

Methods	**PFS**			**OS**		
	C-index(↑)	1y AUC(↑)	2y AUC(↑)	C-index(↑)	1y AUC(↑)	2y AUC(↑)
IPI	$0.714_{0.007}$	$0.755_{0.005}$	$0.758_{0.004}$	$0.749_{0.014}$	$0.722_{0.025}$	$0.785_{0.016}$
Ann Arbor	$0.632_{0.001}$	$0.667_{0.000}$	$0.669_{0.007}$	$0.667_{0.003}$	$0.665_{0.016}$	$0.680_{0.005}$
Meta. Param.†	$0.705_{0.006}$	$0.734_{0.012}$	$0.736_{0.007}$	$0.766_{0.018}$	$0.791_{0.046}$	$0.795_{0.018}$
Meta. Param.*	$0.724_{0.004}$	$0.766_{0.003}$	$0.756_{0.008}$	$0.787_{0.006}$	$\mathbf{0.825}_{0.021}$	$0.830_{0.012}$
Radiomics [18]	$0.754_{0.006}$	$0.800_{0.001}$	$0.792_{0.011}$	$0.771_{0.014}$	$0.750_{0.032}$	$0.806_{0.018}$
Guo2021 [25]	$0.778_{0.004}$	$0.816_{0.002}$	$0.801_{0.008}$	$0.801_{0.002}$	$0.790_{0.007}$	$0.822_{0.003}$
Liu2022 [27]	$0.719_{0.001}$	$0.718_{0.011}$	$0.708_{0.002}$	$0.797_{0.003}$	$0.807_{0.009}$	$0.807_{0.002}$
Thiery2025 [29]	$0.758_{0.003}$	$0.839_{0.010}$	$0.782_{0.005}$	$0.809_{0.001}$	$0.820_{0.001}$	$0.843_{0.002}$
LAMP (ours)	$\mathbf{0.791}_{0.002}$	$\mathbf{0.844}_{0.002}$	$\mathbf{0.823}_{0.004}$	$\mathbf{0.828}_{0.001}$	$0.782_{0.013}$	$\mathbf{0.863}_{0.004}$

* 1y AUC and 2y AUC refer to the Time-AUC at the 1-year and 2-year time points. The metabolic parameters are computed based on manual † and automatic segmentation*. The larger numbers within each cell in the table represent the mean values, while the smaller numbers denote the standard deviations.

Two metrics, the time-dependent area under the receiver operating characteristic curve (Time-AUC) and the concordance index (C-index), are used to evaluate the prognosis performance. To ensure data integrity and prevent leakage, we applied the same five-fold cross-validation split for both the segmentation network and the prognosis network training. To extract feature maps with the same spatial dimensions as the input images, we implemented patch-wise training and inference during the segmentation network training.

Method Comparison. To evaluate the performance of the proposed model, we compared it against four categories of baselines: (1) Widely-used clinical prognosis indices, including IPI and Ann Arbor staging (Ann Arbor) [10,11]; (2) Metabolic parameters (Meta. Param.) [34] extracted from segmentation masks, including those derived from both manual and automatic segmentation methods; (3) End-to-end whole image-based models, such as Guo2021 [25],Liu2022 [27]; and (4) Graph Convolutional Networks (GCNs) [29]. As shown in Table 1, the performance of the proposed model outperforms the currently used prognostic clinical indices and metabolic parameters. The C-index for the proposed model shows an improvement of over 9% and 5% for PFS and OS, respectively, compared to the best-performing model. This demonstrates the superior predictive power of our model in both endpoints. For Guo2021 [25] and Liu2022 [27], the constraint of a fixed input size may hinder model performance by limiting its ability to capture variations in lesion size and distribution. Moreover, jointly performing segmentation and prognosis prediction within a single model [27] introduces challenges in task balancing, potentially leading to suboptimal learning

Table 2. Ablation studies under five-fold cross-validation.

Methods	**PFS**			**OS**		
	C-index(↑)	1y AUC(↑)	2y AUC(↑)	C-index(↑)	1y AUC(↑)	2y AUC(↑)
Baseline	$0.750_{0.003}$	$0.816_{0.004}$	$0.775_{0.003}$	$0.794_{0.008}$	$\mathbf{0.844}_{0.015}$	$0.822_{0.011}$
w/ ABMLA	$0.771_{0.002}$	$0.802_{0.005}$	$0.777_{0.003}$	$0.800_{0.003}$	$0.724_{0.016}$	$0.827_{0.006}$
w/ LACF	$0.765_{0.001}$	$0.841_{0.002}$	$0.796_{0.001}$	$0.813_{0.008}$	$0.795_{0.028}$	$0.842_{0.011}$
Full model	$\mathbf{0.791}_{0.002}$	$\mathbf{0.844}_{0.002}$	$\mathbf{0.823}_{0.004}$	$\mathbf{0.828}_{0.001}$	$0.782_{0.013}$	$\mathbf{0.863}_{0.004}$

* The larger numbers within each cell in the table represent the mean values, while the smaller numbers denote the standard deviations.

of both objectives. While Thiery2025 [29] incorporate inter-lesional propagation relationships, GCN lacks the integration of perilesional anatomical context, which may contribute to unfavorable performance. The last, the Kaplan-Meier curves presented in Fig. 2. A and 2. B demonstrates a statistically significant distinction ($p \leq 0.05$) between high-risk and low-risk patients. This finding further validates the model's discriminative capacity and underscores its potential clinical translational value.

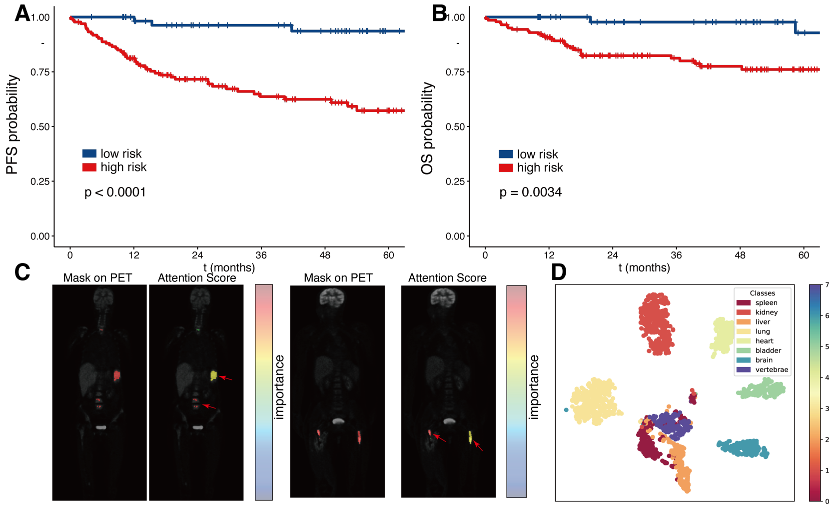

Fig. 2. (A) and (B) are the Kaplan-Meier (KM) curves for PFS and OS prognosis. (C) presents mapping the attention scores $a^{i}_{i_j}$ over the lesion regions. (D) shows the features of key anatomical organs by t-SNE.

Ablation Study. Table 2 highlights two key advantages of our proposed model: ABMLA and LACF, both jointly contribute to performance improvement compared with baselines (all features are aggregated by averaging). Additionally, in Fig. 2. C, we visualized the attention scores learned through ABMLA. A higher attention score is assigned to the lesion regions within the bone marrow and spleen, as involvements of these organs typically indicate a poor prognosis. The t-SNE scatter map in Fig. 2. D indicates that features extracted from the segmentation network can differentiate key organs, which is beneficial for the prognosis task via our LACF strategy.

4 Conclusion

For malignancies like lymphoma, which have multiple, variable lesion regions and challenges in feature representation and aggregation, we propose a PET/CT-based model LAMP that combines attention-based multi-lesion aggregation with contextual interaction with surrounding anatomical structures. Our approach is comprehensively evaluated on PFS and OS prognosis using multiple metrics. The results demonstrate the superiority of our model both quantitatively and qualitatively, highlighting its potential to serve as an auxiliary tool to assist clinicians in treatment planning.

Acknowledgments. This work was supported by DAMO Academy through DAMO Academy Research Intern Program.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Sehn, L. H., Salles, G.: Diffuse large B-cell lymphoma. New England Journal of Medicine **384**(9), 842–858 (2021)
2. Li, S., et al.: Diffuse large B-cell lymphoma. Pathology **50**(1), 74–87 (2018)
3. Siegel, R. L., et al.: Cancer statistics, 2024. CA: a cancer journal for clinicians **74**(1), 12–49 (2024)
4. Susanibar-Adaniya, S., Barta, S. K.: 2021 update on diffuse large B cell lymphoma: a review of current data and potential applications on risk stratification and management. American Journal of Hematology **96**(5), 617–629 (2021)
5. Tilly, H., et al.: Diffuse large B-cell lymphoma (DLBCL): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology **26**, v116–v125 (2015)
6. Roschewski, M.,et al.: Diffuse large B-cell lymphoma—treatment approaches in the molecular era. Nature Reviews Clinical Oncology **11**(1), 12–23 (2014)
7. Eertink, J. J., et al.: Optimal timing and criteria of interim PET in DLBCL: a comparative study of 1692 patients. Blood Advances **5**(9), 2375–2384 (2021)
8. Cabero Martinez, A., et al.: Early Identification of Very High Risk Patients with Diffuse Large B Cell Lymphoma (DLBCL) By PET-CT. Blood **142**(Supplement 1), 4511–4511 (2023)
9. Wenzl, K., et al.: Multiomic analysis identifies a high-risk signature that predicts early clinical failure in DLBCL. Blood Cancer Journal **14**(1), 100 (2024)
10. Shipp, M. A.: International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. New England Journal of Medicine **329**, 987–994 (1993)
11. Ruppert, A. S., et al.:International prognostic indices in diffuse large B-cell lymphoma: a comparison of IPI, R-IPI, and NCCN-IPI. Blood **135**(23), 2041–2048 (2020)
12. Wang, Runze and Zheng, Guoyan: CyCMIS: Cycle-consistent Cross-domain Medical Image Segmentation via diverse image augmentation, Medical Image Analysis (76), 102328, (2022)
13. Wang, Runze, et al.:Few-shot medical image segmentation regularized with self-reference and contrastive learning, International Conference on Medical Image Computing and Computer-Assisted Intervention, 514-523, (2022)
14. Ceriani, L., et al.:Metabolic heterogeneity on baseline 18FDG-PET/CT scan is a predictor of outcome in primary mediastinal B-cell lymphoma. Blood **132**(2), 179–186 (2018)
15. Sasanelli, M., et al.:Pretherapy metabolic tumour volume is an independent predictor of outcome in patients with diffuse large B-cell lymphoma. European Journal of Nuclear Medicine and Molecular Imaging **41**, 2017–2022 (2014)
16. Mikhaeel, N. G., et al.:Combination of baseline metabolic tumour volume and early response on PET/CT improves progression-free survival prediction in DLBCL. European Journal of Nuclear Medicine and Molecular Imaging **43**, 1209–1219 (2016)
17. Esfahani, S. A., et al.:Baseline total lesion glycolysis measured with 18F-FDG PET/CT as a predictor of progression-free survival in diffuse large B-cell lymphoma: a pilot study. American Journal of Nuclear Medicine and Molecular Imaging **3**(3), 272 (2013)
18. Carlier, T., et al.: Prognostic value of 18F-FDG PET Radiomics features at baseline in PET-Guided consolidation strategy in diffuse large B-Cell lymphoma: a machine-learning analysis from the GAINED study. Journal of Nuclear Medicine **65**(1), 156–162 (2024)

19. Eertink, J. J.,et al.: 18F-FDG PET/CT baseline radiomics features improve the prediction of treatment outcome in diffuse large B-cell lymphoma patients. Blood **136**, 27–28 (2020)
20. Eertink, J. J., et al.: Baseline PET radiomics outperforms the IPI risk score for prediction of outcome in diffuse large B-cell lymphoma. Blood **141**(25), 3055–3064 (2023)
21. Zhong, Y., et al.:PET/CT based cross-modal deep learning signature to predict occult nodal metastasis in lung cancer. Nature Communications **14**(1), 7513 (2023)
22. Meng, M., et al.: DeepMTS: Deep multi-task learning for survival prediction in patients with advanced nasopharyngeal carcinoma using pretreatment PET/CT. IEEE Journal of Biomedical and Health Informatics **26**(9), 4497–4507 (2022)
23. Cheng, N. M., et al.: Deep learning for fully automated prediction of overall survival in patients with oropharyngeal cancer using FDG-PET imaging. Clinical Cancer Research **27**(14), 3948–3959 (2021)
24. Häggström, I., et al.: Deep learning for [18F] fluorodeoxyglucose-PET-CT classification in patients with lymphoma: a dual-centre retrospective analysis. The Lancet Digital Health **6**(2), e114–e125 (2024)
25. Guo, R., et al.: Weakly supervised deep learning for determining the prognostic value of 18 F-FDG PET/CT in extranodal natural killer/T cell lymphoma, nasal type. European Journal of Nuclear Medicine and Molecular Imaging, 1–11 (2021)
26. Yuan, C., et al.: Multimodal deep learning model on interim [18F] FDG PET/CT for predicting primary treatment failure in diffuse large B-cell lymphoma. European Radiology **33**(1), 77–88 (2023)
27. Liu, P., et al.: Joint lymphoma lesion segmentation and prognosis prediction from baseline FDG-PET images via multitask convolutional neural networks. IEEE Access **10**, 81612–81623 (2022)
28. Lv, W., et al.: Functional-structural sub-region graph convolutional network (FS-GCN): application to the prognosis of head and neck cancer with PET/CT imaging. Computer Methods and Programs in Biomedicine **230**, 107341 (2023)
29. Thiery, O., et al.: PET-based lesion graphs meet clinical data: An interpretable cross-attention framework for DLBCL treatment response prediction. Computerized Medical Imaging and Graphics **120**, 102481 (2025)
30. Yang, S., et al.: A Matched Comparison of Pola-R-CHP Vs. R-CHOP for DLBCL with Extranodal Involvement. Blood **144**, 6478 (2024)
31. Isensee, F., et al.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods **18**(2), 203–211 (2021)
32. Wasserthal, J., et al.: TotalSegmentator: robust segmentation of 104 anatomic structures in CT images. Radiology: Artificial Intelligence **5**(5), e230024 (2023)
33. Chen, R. J., et al.: Multimodal co-attention transformer for survival prediction in gigapixel whole slide images. Proceedings of the IEEE/CVF International Conference on Computer Vision 2021, pp. 1–13 (2021)
34. Frood, R., et al.:Baseline PET/CT imaging parameters for prediction of treatment outcome in Hodgkin and diffuse large B cell lymphoma: a systematic review. European Journal of Nuclear Medicine and Molecular Imaging **48**, 3198–3220 (2021)
35. Huang, Wei, et al.: LIDIA: Precise Liver Tumor Diagnosis on Multi-Phase Contrast-Enhanced CT via Iterative Fusion and Asymmetric Contrastive Learning, International Conference on Medical Image Computing and Computer-Assisted Intervention, 394-404, (2024).
36. Li, Chunli, et al.: Improved Esophageal Varices Assessment from Non-contrast CT Scans, International Conference on Medical Image Computing and Computer-Assisted Intervention, 349-359, (2024).