The Difference Between Radiomic Print of Hodgkin and Non-Hodgkin Lymphoma

Elsayed WA[1], Muhammed A[2], Elnakib E[1], Khalil M[3], Fathy Z[3], Shehata SE[4], Diab WA[4]

[1] Nuclear Medicine Department, Sohag University Hospital, Egypt
[2] Clinical Oncology Department, Sohag University Hospital, Egypt
[3] Radiotherapy and Nuclear Medicine Department- South Egypt Cancer Institute, Assiut University, Egypt
[4] Clinical Oncology and Nuclear Medicine Department Assiut University, Egypt

Received: 29 October 2024
Accepted: 17 November 2024

Abstract:

Background: Lymphoma is one of the leading causes of death in adults and children. Diagnosing a specific subclass of lymphoma requires comprehensive histological evaluation and immune histochemistry analysis. In this study, we examine the ability of an artificial intelligence-based classifier to differentiate between Hodgkin disease (HD) and non-Hodgkin lymphoma (NHL) based on their radiomic print.

Methods: A retrospective cohort was conducted for the patients diagnosed with lymphoma between 2019 and 2023. The initial baseline PETCT scans of these patients were retrieved. The active lesions were segmented and the radiomic features were extracted. The collected features were split into a training set (80%) and a validation set (20%). The primary endpoint of this study was used to build a classifier that could predict the type of lymphoma (Hodgkin or Non-Hodgkin). The training set was used to develop the model and the validation set was used to validate the results.

Results: The study included 78 patients. Hodgkin disease was seen in 51 patients. The total number of identified and segmented lesions was 222, and 111 of them were retrieved from HD scans. Radiomic features were extracted from the PETCT. Several modelling approaches were examined. The highest accuracy was seen with the TabPFN classifier with a validation set accuracy of 73.3%. The model achieved an F1-score of 0.76 and 0.70 for HD and NHL, respectively.

Conclusion: The TabPFN-based classifier achieved an accuracy of 73.3% on the validation sets. Further research on large sets is necessary.

Keywords: Radiomics, radiomics in lymphoma, difference between Hodgkin and non-Hodgkin lymphoma, AI in lymphoma.

Authors Information:

Wafaa Abd-Elhamied Elsayed
Nuclear Medicine Department, Sohag University Hospital, Egypt.
email: wafaamagraby@yahoo.com

Amr Muhammed
Clinical Oncology Department, Sohag University Hospital, Egypt.
email: Amr.muhammed@med.sohag.edu.eg

Esraa Elnakib
Nuclear Medicine Department, Sohag University Hospital, Egypt.
email: Dr.esraa.hafez@gmail.com

Maha Khalil
Radiotherapy and Nuclear Medicine Department- South Egypt Cancer Institute, Assiut University.
email: Mahakhalil12@yahoo.com

Zainab Fathy
Radiotherapy and Nuclear Medicine Department- South Egypt Cancer Institute, Assiut University.
email: Zeinab.fathy.95@gmail.com

Samir Eid Shehata
Clinical Oncology and Nuclear Medicine Department Assiut University.
email: Samir_eid@hotmail.com

Waleed A Diab
Clinical Oncology and Nuclear Medicine Department Assiut University.
email: Waleed.diab@aun.edu.eg

Corresponding Author:
Wafaa Abd-Elhamied Elsayed
Nuclear Medicine Department, Sohag University Hospital, Egypt.
email: wafaamagraby@yahoo.com

Background:

Lymphoma represents is ranked within the top ten types of cancer that affect adults and paediatric populations. Since the discovery of the disease in 1832, there have been enormous efforts to characterise lymphoma into different subclasses [1]. Each of these subsets mandates clinical application of specialised management protocols. Classically, several steps are necessary to achieve correct and robust histopathological diagnosis. They include fixation of the pathological sample followed by embedding and proper sectioning. Finally, the samples are stained by

haematoxylin and eosin, and specialised monoclonal antibodies that target specific antigens within the malignant cells. Afterwards, experienced pathologists would review the results and confirm the final diagnosis. Naturally, this step takes between ten and fifteen days depending on the experiences and equipment within each of the laboratory [2].

Radiomics focuses on automated or semi-automated quantitative feature extraction from medical imaging. It converts biomedical images into a high-dimensional matrix which represents the hidden features within these images. These features can be processed quantitatively by advanced computer algorithms. The output of these processing steps would help in the qualitative evaluation of the tumours [3]. The classic workflow of radiomics extractions depends on image accusations, segmentation of the lesions of interest, feature extractions by specialised protocols and finally data analysis [4]. What differentiates radiomics from conventional qualitative radiology is its dependence on reproducible quantitative measurements rather than human-eye evaluation. This would offer a stable and reliable interpretation of clinical data within the radiographic scans [5, 6].

Several authors examined the role of radiomics in lymphomas. These projects mainly focused on predicting outcomes post-therapy in specific subsets of the disease such as Mantel cell lymphoma and Diffuse large B-cell lymphoma [7, 8]. In our project, we examine the possibility of differentiating Hodgkin lymphoma from non-Hodgkin lymphoma based on their radiomic features and explore the possibility of using this in clinical settings.

Patients and Methods:

Trial Design and Conduct

We conducted a retrospective review of PETCT scans of patients who had been diagnosed with lymphoma between 2019 and 2023. The primary aim of this study is to build a machine-learning model to predict the histological subtype of lymphoma (Hodgkin or Non-Hodgkin) based on its radiomic print in the initial PETCT scan.

Patients' Population

The patient population included patients who were diagnosed with lymphoma (HD and NHL) and had PETCT scans for their initial staging. The DICOM data of these scans were retrieved and included in this study.

Study Endpoint

The primary endpoint was to examine the ability of machine learning to classify the lymphoma into Hodgkin or non-Hodgkin based on their radiomic print.

Machine Learning

The PETCT DICOM were explored by Slicer version 5.4.0 where positive lesions were manually contoured and used for extraction of radiomic features. Inter-rater variability in lesion contours was minimized by employing dual manual segmentation from two experienced nuclear physicians. The First-order statistics, Grey Level Co-occurrence Matrix (GLCM), Gray Level Dependence Matrix (GLDM), Gray Level Run Length Matrix (GLRLM), Gray Level Size Zone (GLSZM), Neighbouring Gray Tone Difference Matrix (NGTDM), Shape-based 3D, and Shape-based 2D feature classes were used to generate the radiomic features.

The extracted features were manipulated later by the Python 3.10 package, NumPy and Pandas. Several steps were taken to achieve data preprocessing. They included conversion of data to float points 'float32', class imbalance correction by RandomOverSampler of the Imblearn library, individual feature rescaling between 0 and 1.0, and dimensionality reduction by Principal Component Analysis (PCA) of Sci-kit learn. The step of PCA was necessary to reduce the features to 20 and meet the threshold of the TabPFN classifier. The final step of preprocessing was randomly splitting the data into training and test sets with a split size of 20% for the test set. Afterwards, the TabPFN library was used to build a classification model with 16 ensemble configurations. The process is summarised in Figure 1. The performance of the proposed model was compared to several other machine learning approaches including XGboost, random forest classifier, Bagging classifier, and Feed-forward neural network.

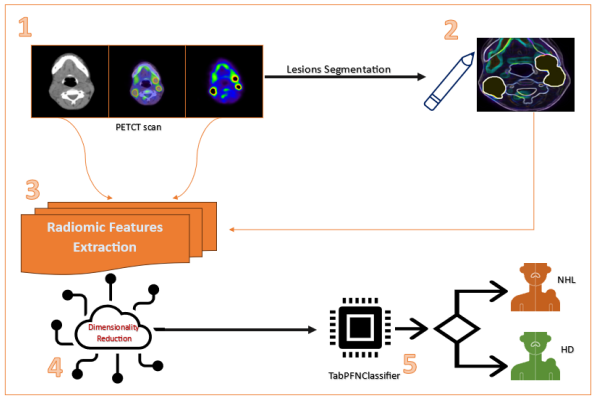

Figure 1 – The Process of radiomics extractions and data processing.

1. The PETCT data were retrieved from the DICOM server.
2. The disease sites were segmented.
3. The contoured volumes were processed by Slicer for radiomic feature extraction.
4. The Extracted Features were concatenated, preprocessed (Imbalance correction and Min/Max Scaling), and had dimensionally reducted.
5. The data was processed by the TabPFN classifier for disease classification.

Ethics statement

The Research Ethics Board of the hospital has approved the study. Informed consents were obtained from the participants or their legal guardians. Unique codes were used to de-identify the patients' names and maintain confidentiality.

Results:

Patient demographics and flow of data processing

The total number of patients included in this study was 78. Hodgkin disease was seen in 51 of them. The utter number of extracted lesions was 222. HD disease compromised 111 of them. Table 1 illustrates the anatomical distri4bution of the included lesions.

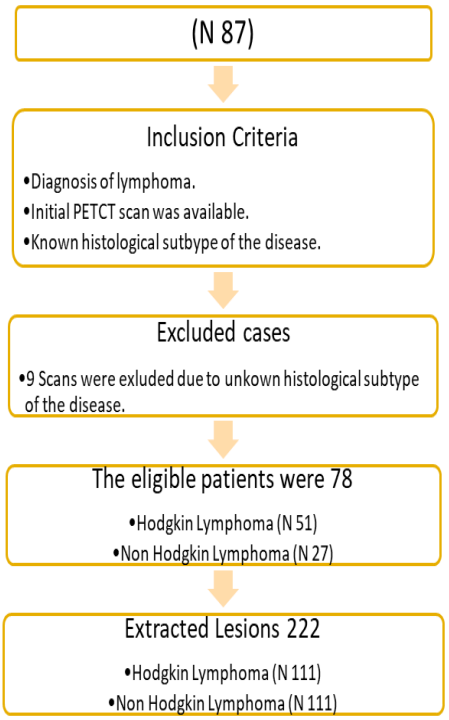

Figure 2 outlines the sequential stages involved in the data collection process.

Table 1 - Anatomical Distribution of included lesion

	HD (n = 111)	NHL (n = 111)
Cervical Lymph Nodes	57	45
Mediastinal Lymph Node	28	39
Para-aortic Lymph Nodes	15	21
Iliac Nodes Lymph Node	6	4
Spleen	5	2

TabFPN Performance

The model achieved an accuracy of 97.7% on the training set and 73.3% on the test set. The precision, recall and F1 scores of the model were included in Table 2. Figures 3 and 4 show the confusion matrix of the training and test sets, respectively.

Table 3 illustrates the performance among other comparative models. These comparative models were XG-Boost, Random Forest Classifier (RFC), Bagging Classifier, and Feeding Forward neural network (FFN). The reported models either suffered from broad underperformance or overfitting the training set. However, the best performing model was RFC with accuracy of 0.85 and 0.67 for the training and validation set, respectively. The worst performance was seen by FFN with accuracy of 0.50, and 0.44 for training and validation, respectively.

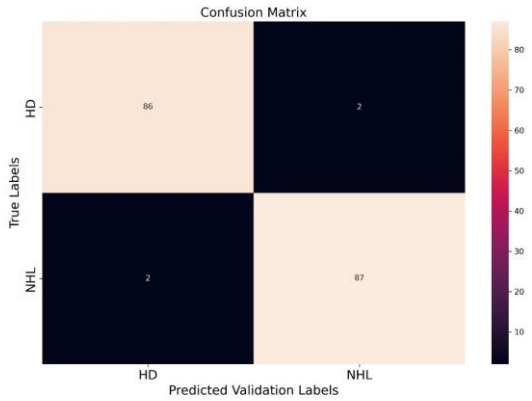

Figure 3 - Confusion matrix of the training set showing accuracy of prediction of the two different subclasses.

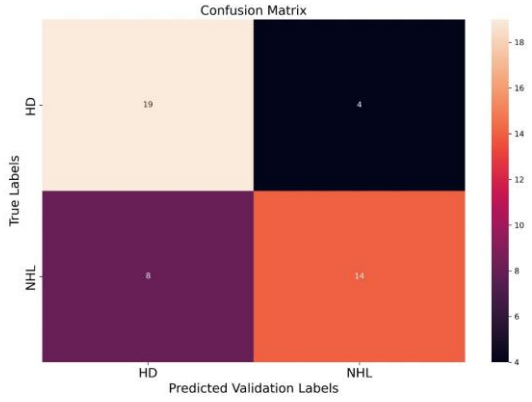

Figure 4 - Confusion matrix of the validation set showing accuracy of prediction of the two different subclasses.

Table 2 – Precision, Recall and F1 Score of the proposed model

	Precision		Recall		F-1 Score		
	HD	NHL	HD	NHL	Total	HD	NHL
Training Set	0.98	0.98	0.98	0.98	0.98	0.98	0.98
Test Set	0.70	0.78	0.83	0.64	0.73	0.76	0.70

Table 3 – Performance of other comparative models

		XGB Boost[1]	RFC[2]	Bagging Classifier[3]	Neural Network[4]
Accuracy	Training	1.0	0.85	0.63	0.50
	Validation	0.64	0.67	0.51	0.44
F1 Score	Training	1.0	0.85	0.56	0.25
	Validation	0.66	0.68	0.42	0.13

1- *XGB boost: max depth of 3, number of estimators 150, and minimum child weight 3.*
2- *Random Forest Classifier: max depth of 3, number of estimators 100.*
3- *Bagging Classifier: Logistic regression as base estimator, number of estimators 20.*
4- *Neural network: three dense layers (512, 512, 256), each followed by ReLU activation layer, Batch-Normalization and Dropout of 0.3. Final layer had single node with Sigmoid Activation function.*

Discussion:

PETCT was introduced in 1998 as functional imaging that expresses the metabolic signature of tissues over the anatomical landmarks [9]. Several traditional approaches such as metabolic tumour volume (MTV), total lesion glycolysis (TLG), intra-tumoral heterogeneity, and standardised uptake value (SUV) were proposed to characterise the found lesions inside the PETCT scans. These parameters reflected the activity within the tumours. Also, they were linked to response to treatment or worse prognosis [3, 10].

The clinical benefit from these parameters makes it clear that the quantitative evaluation of PETCT is worth further exploration. Therefore, radiomics was introduced in this setting [11-13]. Radiomics relies on mathematical methods that can define lesions quantitatively. It focuses on describing the size, shape, morphology, heterogeneity of texture and edge of the lesions. Simply, Radiomics features are just matrix representations of the lesions of interest within an image. This extracted information can be processed further to conclude and test different clinical hypotheses [14].

Lymphoma as a disease exhibited the largest footstep in the field of radiomics. Earlier studies showed the ability of radiomics to discriminate lymphomas from non-lymphomas [11, 15]. Moreover, several studies found that radiomics can correlated with lymphomatous infiltration of the bone marrow, and disease prognosis [8, 16]. Most of these studies just focused on a single subset of lymphoma such as Hodgkin's disease, diffuse B-cell lymphoma or mantel cell lymphoma. Also, these studies varied in the number of features included in their analysis. Some of them adapted a few parameters such as metabolic bulk volumes, or heterogeneity index, while others included hundreds of features in their methodology [12].

To our knowledge, only one study tried to use extracted features from the CT scan of the PETCT to differentiate between different subsets of lymphoma. The authors included 71 patients, and nearly two-thirds of them had non-Hodgkin lymphoma. They used multilayer perceptron and achieved an accuracy of 75.76% in their validation set [17]. In our model, examined 78 patients. Roughly, two-thirds of them had Hodgkin's disease. We used Slicer's segment editor to do manual segmentation. Then, we extracted the radiomic features from both the CT and PET images. Our TabPFN classifier achieved an accuracy of 97.7% on the training set and 73.3% on the test set. Also, it outperformed the other proposed model architecture in Table 3.

Clinical Significance of the model

While clinical guidelines emphasize the importance of histopathological classification for lymphoma diagnosis before treatment, this isn't always clinically feasible, especially in urgent situations. The traditional immune histochemistry analysis can take several days due to tissue processing and interpretation by

pathologists [2]. In life-threatening presentations such as spinal cord compression or high tumour burden leading to multi-organ failure, such delays can have serious life-threatening consequences [18, 19]. We understand that our model is by far less accurate compared to the conventional approaches in classifying the disease, but it can compete effectively when it comes to the time perspective. This speed may be valuable to haematologists and oncologists in expediting the initiation of potentially life-saving therapies in time-sensitive situations, while concurrently laying the groundwork for more definitive diagnoses and targeted treatment plans once conventional methods become available.

Moreover, our model could potentially facilitate rapid screening of all nodal groups within the body, offering insights into the class of lymphoma. This could contribute to addressing limitations inherent in the typical diagnostic approach, which often relies solely on histopathological analysis of a single excised lymph node [20]. While convenient, this approach carries an inherent limitation: it assumes the excised node reflects the entire disease burden throughout the body, potentially overlooking the presence of concurrent Hodgkin and non-Hodgkin lymphomas, collectively known as Poly-lymphomatous Syndrome (PLS) [21-23]. The precise incidence of PLS remains unknown, with documented case series and reports suggesting its existence. Estimates suggest it might be more prevalent than currently recognized [23]. In some cases, the presence of PLS is suspected when patients exhibit poor treatment response, prompting additional biopsies from non-responsive sites to investigate this possibility [24, 25]. Screened nodes exhibiting discrepancy from the expected disease type should undergo additional biopsy and evaluation. This approach would lead to pre-emptive diagnosis of conditions such as PLS instead of waiting several weeks before finding a patient is not responding well to cytotoxic therapies.

Furthermore, expanding research on lymphoma characterization using radiomics could offer valuable insights into potential cost-effectiveness gains compared to conventional methods. While the average cost of performing immune histochemistry analysis on a single sample ranges between 500 and 1000 USD [26, 27]. On the other hand, the global estimate of approximately one million new lymphoma cases yearly highlights the potential impact of improving diagnostic accuracy and streamlining workflows [28]. Therefore, collaborative efforts to build large datasets of diverse PET-CT scans represent a promising avenue for developing more robust classifiers. However, directly attributing significant cost reductions to these classifiers compared to conventional methods remains premature and requires comprehensive analysis beyond the scope of this research.

Limitation

Our study has two major limitations. The first was the retrospective nature of the data making it less prone to bias. Also, the included sample size was 78 patients constraining its power significantly.

Conclusion:

We developed a model based on the TabPFN classifier, which supports previous findings demonstrating the ability of radiomics to characterize lesions in PET-CT scans. Our model achieved an accuracy of 73.3% in predicting the specific subtype of lymphoma on the validation set. Further research is necessary, particularly as such technologies have the potential to improve cost-effectiveness, especially in resource-constrained healthcare systems.

Conflict of interest

The authors declare no conflict of interest.

Authors' contributions

SE, WD contributed to the conception and design of the work, AM, WA, EE, MK, ZF contributed to the acquisition, analysis, and interpretation of the data. AM, WA, EE, ZF wrote the initial draft of the manuscript. SE, WA, MK, EE, AM revised and supervised the work. All authors approved the final version of the manuscript.

Acknowledgement

The authors have no acknowledgement regarding this manuscript.

References:

1. Lakhtakia R, Burney I. A Historical Tale of Two Lymphomas: Part I: Hodgkin lymphoma. Sultan Qaboos Univ Med J. 2015 May;15(2):e202-6.
2. Fox CP, Chaganti S, McIlroy G, et al. The management of newly diagnosed large B-cell lymphoma: A British Society for Haematology Guideline. Br J Haematol. 2024 Apr;204(4):1178-1192.
3. Zhou Y, Zhu Y, Chen Z, et al. Radiomic Features of (18)F-FDG PET in Hodgkin Lymphoma Are Predictive of Outcomes. Contrast Media Mol Imaging. 2021 Nov 22;2021:6347404.
4. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012 Mar;48(4):441-6.
5. Leijenaar RT, Carvalho S, Velazquez ER, et al. Stability of FDG-PET Radiomics features: an integrated analysis of test-retest and inter-observer variability. Acta Oncol. 2013 Oct;52(7):1391-7.
6. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2016 Feb;278(2):563-77.
7. Mayerhoefer ME, Riedl CC, Kumar A, et al. Radiomic features of glucose metabolism enable prediction of outcome in mantle cell lymphoma. Eur J Nucl Med Mol Imaging. 2019 Dec;46(13):2760-2769..
8. Cottereau AS, Nioche C, Dirand AS, et al. (18)F-FDG PET Dissemination Features in Diffuse Large B-Cell Lymphoma Are Predictive of Outcome. J Nucl Med. 2020 Jan;61(1):40-45.

9. Jones T, Townsend D. History and future technical innovation in positron emission tomography. J Med Imaging (Bellingham). 2017 Jan;4(1):011013.
10. Guo B, Tan X, Ke Q, et al. Prognostic value of baseline metabolic tumor volume and total lesion glycolysis in patients with lymphoma: A meta-analysis. PLoS One. 2019 Jan 9;14(1):e0210224.
11. Ben Bouallègue F, Tabaa YA, Kafrouni M, et al. Association between textural and morphological tumor indices on baseline PET-CT and early metabolic response on interim PET-CT in bulky malignant lymphomas. Med Phys. 2017 Sep;44(9):4608-4619.
12. Chauvie S, Ceriani l, Zucca E, Radiomics, in Malignant Lymphoma. 2021, Exon Publications.
13. Cheng JA, Lin YC, Lin Y, et al. Machine Learning Radiomics Signature for Differentiating Lymphoma versus Benign Splenomegaly on CT. Diagnostics (Basel). 2023 Dec 8;13(24):3632.
14. Sollini M, Cozzi L, Antunovic L, et al. PET Radiomics in NSCLC: state of the art and a proposal for harmonization of methodology. Sci Rep. 2017 Mar 23;7(1):358.
15. Dong C, Zheng YM, Li J, et al. A CT-based radiomics nomogram for differentiation of squamous cell carcinoma and non-Hodgkin's lymphoma of the palatine tonsil. Eur Radiol. 2022 Jan;32(1):243-253..
16. Parvez A, Tau N, Hussey D, et al. (18)F-FDG PET/CT metabolic tumor parameters and radiomics features in aggressive non-Hodgkin's lymphoma as predictors of treatment outcome and survival. Ann Nucl Med. 2018 Jul;32(6):410-416.
17. Erturk H, Eser MB, Yaşar AB, et al. Low-dose CT radiomics features-based neural networks predict lymphoma types. Egyptian Journal of Radiology and Nuclear Medicine 2023; 54(1):135.
18. Ghedira K, Matar N, Bouali S, et al. Hodgkin Lymphoma revealed by epidural spinal cord compression. J Spinal Cord Med. 2019 May;42(3):402-404.
19. Chua BJG, Low CE, Yau CE, et al. Recent updates on central nervous system prophylaxis in patients with high-risk diffuse large B-cell lymphoma. Exp Hematol Oncol. 2024 Jan 3;13(1):1.
20. Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. J Clin Oncol. 2014 Sep 20;32(27):3059-68.
21. Bakhshi S, Vidushi A, Das P. Simultaneous Occurrence of Hodgkin Lymphoma and Non-Hodgkin Lymphoma in Siblings. J Pediatr Hematol Oncol. 2009 Nov;31(11):888-90.
22. Al-Mansour M, Connors JM, Gascoyne RD, et al. Transformation to aggressive lymphoma in nodular lymphocyte-predominant Hodgkin's lymphoma. J Clin Oncol. 2010 Feb 10;28(5):793-9.
23. Iqbal M, Jiang L, Li KD, et al. Poly-lymphomatous Syndrome With Concurrent or Sequential Hodgkin and Non-Hodgkin lymphoma. Clin Lymphoma Myeloma Leuk. 2023 Feb;23(2):138-144.
24. Hartmann S, Eray M, Döring C, et al. Diffuse large B cell lymphoma derived from nodular lymphocyte predominant Hodgkin lymphoma presents with variable histopathology. BMC Cancer. 2014 May 13;14:332.
25. Balodis A, Pimenova A, Nikulshin S, et al. Rare Case of Hodgkin Lymphoma Transformation into Diffuse Large B-Cell Lymphoma with Atypical Spread Epidurally, Intradurally and Intramedullary: A Case Report. Am J Case Rep. 2022 Feb 25;23:e935014.
26. Raab SS. The Cost-Effectiveness of Immunohistochemistry. Arch Pathol Lab Med. 2000 Aug;124(8):1185-91.
27. Valvert F, Silva O, Solórzano-Ortiz E, et al. Low-cost transcriptional diagnostic to accurately categorize lymphomas in low- and middle-income countries. Blood Adv. 2021 May 25;5(10):2447-2455.
28. Chu Y, Liu Y, Fang X, et al. The epidemiological patterns of non-Hodgkin lymphoma: global estimates of disease burden, risk factors, and temporal trends. Front Oncol. 2023 Jun 2;13:1059914.