Am J Cancer Res 2023;13(2):509-525
www.ajcr.us /ISSN:2156-6976/ajcr0146852

Original Article

Prognostic value of genetic alterations and ^{18}F-FDG PET/CT imaging features in diffuse large B cell lymphoma

Blanca Ferrer-Lores[1], Jose Lozano[2], Almudena Fuster-Matanzo[2], Irene Mayorga-Ruiz[2], Paula Moreno-Ruiz[2], Fuensanta Bellvís[2], Ana B Teruel[5], Ana Saus[1], Alfonso Ortiz[1], Eva Villamón-Ribate[1], Alicia Serrano-Alcalá[1], José L Piñana[1], Pablo Sopena[3], Rosa Dosdá[4], Carlos Solano[5], Ángel Alberich-Bayarri[2], María José Terol[1,5]

[1]Hematology Department, Hospital Clínico Universitario-INCLIVA, Valencia, Spain; [2]Quantitative Imaging Biomarkers in Medicine, Quibim, Valencia, Spain; [3]Nuclear Medicine Department, Área Clínica de Imagen Médica, La Fe Hospital, Valencia, Spain; [4]Department of Radiology, Hospital Clínico Universitario Valencia, Valencia, Spain; [5]University of Valencia, Valencia, Spain

Received October 5, 2022; Accepted December 3, 2022; Epub February 15, 2023; Published February 28, 2023

Abstract: The current standard front-line therapy for patients with diffuse large-B cell lymphoma (DLBCL)—rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP)—is found to be ineffective in up to one-third of them. Thus, their early identification is an important step towards testing alternative treatment options. In this retrospective study, we assessed the ability of ^{18}F-FDG PET/CT imaging features (radiomic + PET conventional parameters) plus clinical data, alone or in combination with genomic parameters to predict complete response to first-line treatment. Imaging features were extracted from images prior treatment. Lesions were segmented as a whole to reflect tumor burden. Multivariate logistic regression predictive models for response to first-line treatment trained with clinical and imaging features, or with clinical, imaging, and genomic features were developed. For imaging feature selection, a manual selection approach or a linear discriminant analysis (LDA) for dimensionality reduction were applied. Confusion matrices and performance metrics were obtained to assess model performance. Thirty-three patients (median [range] age, 58 [49–69] years) were included, of whom 23 (69.69%) achieved long-term complete response. Overall, the inclusion of genomic features improved prediction ability. The best performance metrics were obtained with the combined model including genomic data and built applying the LDA method (AUC of 0.904, and 90% of balanced accuracy). The amplification of *BCL6* was found to significantly contribute to explain response to first-line treatment in both manual and LDA models. Among imaging features, radiomic features reflecting lesion distribution heterogeneity (*GLSZM_GrayLevelVariance*, *Sphericity* and *GLCM_Correlation*) were predictors of response in manual models. Interestingly, when the dimensionality reduction was applied, the whole set of imaging features-mostly composed of radiomic features-significantly contributed to explain response to front-line therapy. A nomogram predictive for response to first-line treatment was constructed. In summary, a combination of imaging features, clinical variables and genomic data was able to successfully predict complete response to first-line treatment in DLBCL patients, with the amplification of *BCL6* as the genetic marker retaining the highest predictive value. Additionally, a panel of imaging features may provide important information when predicting treatment response, with lesion dissemination-related radiomic features deserving especial attention.

Keywords: Diffuse large B-cell lymphoma, PET/CT, complete response, predictive models, imaging features, radiomics, genomic alterations, BCL6 amplification, tumor burden, lesion dissemination

Introduction

Diffuse large B-cell lymphoma (DLBCL) is the most common type of non-Hodgkin lymphoma (NHL) both in the United States and Western countries, accounting for around one-third of NHL cases [1, 2]. DLBCL is a highly heterogenous disease at a clinical, pathological, and molecular level, showing different survival outcomes [3]. Despite its aggressive disease course, approximately 50% to 70% of patients may be cured by current standard front-line therapy, an anti-CD-20-based chemoimmunotherapy consisting of rituximab plus cyclophos-

phamide, doxorubicin, vincristine, and prednisone (R-CHOP) [4]. However, R-CHOP is found to be ineffective in up to one-third of patients, due to either primary refractoriness or relapse after reaching a complete response [5-7]. Thus, early identification of patients with poor prognosis is an important step towards testing alternative treatment options.

The International Prognostic Index (IPI), a risk stratification system that encompasses five factors-age >60 years, elevated serum lactate dehydrogenase (LDH), Eastern Cooperative Oncology Group (ECOG) performance status ≥2, Ann Arbor stage ≥III, and >1 extranodal site-remains the primary clinical tool for predicting outcome and for stratifying patients into clinical trials [8, 9]. This index has been validated and refined in the modern rituximab era with the revised IPI (R-IPI) and the National Comprehensive Cancer Network IPI (NCCN-IPI), allowing better discrimination performance [10-12]. However, all three scoring systems fail to identify patients with less than a 50% chance of survival, who are usually patients with primary refractory disease after R-CHOP treatment [9]. Consequently, additional predictors are needed to better characterize this high-risk group of patients requiring new treatment approaches.

On the other hand, molecular aberrations in tumor cells seem to also retain an important prognostic value. Thus, tumors harboring a *MYC* rearrangement concurrent with a rearrangement in B-cell lymphoma 2 (*BCL2*), or B-cell lymphoma 6 (*BCL6*), or both genes (also known as "double hit" and "triple hit" lymphoma, respectively) are associated with transcriptional dysregulation of *MYC* and highly aggressive clinical behavior with resistance to standard chemotherapy and extremely poor outcome [13-18]. In addition to *MYC, BCL2*, and *BCL6* translocations, other less frequent and not so well studied mutations may also affect these genes in DLBCL, such as copy number alterations (CNAs). Although results are controversial, several studies suggest that *MYC* extra copy is an independent poor prognostic factor similar to *MYC* rearrangement [19]. A worse prognosis has been also related with increased *BCL2* copy number [20, 21]. Likewise, patients with both *MYC* and *BCL2* CNAs or with concurrent translocations and copy number gains in *MYC*, *BCL2*, and/or *BCL6* have been reported to show similar outcomes to those with classic "double hit" and "triple hit" lymphoma [22, 23].

At present, positron emission tomography/computed tomography (PET/CT) with fluorodeoxyglucose (^{18}F-FDG) is the standard-of-care imaging modality for patients with DLBCL. It is routinely used for disease staging and treatment response evaluation, allowing personalized therapeutic decision-making [24-27]. The standardized uptake value (SUV), a semiquantitative measurement of uptake in tissue at sites of disease, is the parameter most frequently used in assessment [28]. Along with the total metabolic tumor volume (MTV), this semiquantitative interpretation method of ^{18}F-FDG PET/CT has been reported to have important prognostic and predictive roles, and both have been associated with survival outcomes in patients with DLBCL [29-34]. However, the SUV may be easily affected by multiple factors (e.g. blood glucose level, body weight, scanning protocol, reconstruction parameters, and dose extravasation) and can only provide information on tumor glycolysis [35, 36], while MTV values vary depending on measurement procedures and to date, no standardization exists [37]. Additionally, this parameter does not allow to assess the heterogeneity of the distribution of the lesions, an important factor in DLBCL, which often involves multiple disseminated nodal sites possibly associated with extranodal sites, sometimes with genetic heterogeneity impacting outcome [38]. Consequently, novel PET-derived quantitative imaging biomarkers are required that may further help to individualize lymphoma treatment.

In this sense, radiomics, a high-throughput quantitative imaging analysis method which extracts a large number of features from medical images [39], has emerged as a promising discipline. Indeed, high throughput radiomic features based on texture analysis allow to assess intra-tumoral heterogeneity, a pivotal prognostic factor in cancer progression, recurrence, and therapeutic resistance [40], importantly related to patient outcomes, tumor aggressiveness, metastasis, and molecular profiles [41, 42]. Interestingly, and depending on how lesion segmentation—a critical step in radiomics workflow—is performed, radiomic features could potentially reflect lesion heterogeneity and distribution, both structural and metabolic, an information especially relevant in non-solid tumors such as DLBCL.

Today, the evidence on the predictive value of radiomics in DLBLC is scarce. Some works have reported promising results in the prediction of treatment and/or survival outcomes based on radiomics thanks to machine learning or statistical models, often demonstrating better results when radiomic features are combined with clinical information [43-47]. However, data about the potential of radiomics in combination with genomic information are lacking.

The aim of this study was to assess the ability of ^{18}F-FDG PET/CT imaging features (radiomic features + PET conventional parameters) plus clinical data, alone or in combination with genomic information, all collected at baseline, to predict complete response to first-line treatment in DLBCL patients. This work describes the development of two multivariate logistic regression predictive models for prognosis (model 1: imaging features [radiomic features + PET conventional parameters] + clinical data; model 2: imaging features + clinical data + genomic data) that will be further validated in future studies.

Methods

Study design and patient population

This was a retrospective single-center observational study in accordance with the Declaration of Helsinki and approved by the local ethics committee. No written informed consent from patients was required.

The Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement was followed for this work.

Patients with newly diagnosed DLBCL between January 1st, 2015, and July 31st, 2018 at the University Clinic Hospital of Valencia were retrospectively reviewed. The inclusion criteria were: 1) a diagnosis of aggressive non-Hodgkin lymphoma according to the WHO 2016 [48] including diffuse large B-cell lymphoma and primary mediastinal B-cell lymphoma confirmed by tissue biopsy; 2) available genomic status of *MYC* (8q24.1), *BCL2* (18q21), and *BCL6* (3q27) by FISH on diagnostic biopsy; 3) a ^{18}F-FDG PET/CT scan before treatment; 4) an interval between the PET/CT scan and biopsy of less than 1 month and 5) receiving a first-line chemoimmunotherapy (CHOP-R, CHOP-R-like or R-CODOX-M/R-IVAC-based immunochemotherapy) according to standard clinical guidelines [49, 50]. Patients were excluded if they were lost to follow-up, had incomplete clinical or immunohistochemical data, or had a prior history of solid cancer or any other therapy-related malignancy diagnosed during follow-up.

Clinicopathological data, therapeutic response evaluation and follow-up

Clinical, demographic, and pathological information, Lugano stage, R-IPI stage and histological subtypes were collected from hospital medical records and histologic reports. A complete response (CR) was defined according to Cheson criteria [25] as complete disappearance of all detectable clinical evidence of disease and disease-related symptoms if present before therapy, a score of 1–3 with or without a residual mass based on the Deauville 5-point scale [51, 52] and no evidence of FDG-avid or morphological disease in bone marrow. Long-term CR was defined as CR longer than 3 years. After achieving complete metabolic response, clinical follow-up with clinical history, physical examination, and laboratory work-up was performed every 3 to 4 months for the first 2 years and every 6 months for the following 3 years. From this point onwards, an annual clinical follow-up was performed to assess potential late toxicity and development of second neoplasms.

Genomic alterations

Translocation, amplification or deletion events in *MYC* (8q24.1), *BCL2* (18q21), and *BCL6* (3q27) were determined by FISH break-apart probes (MetaSystems) on Formalin-Fixed Paraffin-Embedded Tissue (FFPE) sections. Briefly, FFPE slides were deparaffinized and then pretreated using Tissue FISH Pretreatment Kit (MetaSystems) following the manufacturer's instructions. Then, 10 µl of probe were added on the samples and the slides were covered with 22 × 22 mm^2 coverslips. Samples were subjected to denaturation (80°C 5 min) and hybridization (37°C overnight) in a programmable hybridizer. Following hybridization, coverslips were removed and the slides were washed in 0.4 × SSC/0.3% NP-40 at 72°C for 2 min and then placed into room temperature 2 × SSC/0.1% NP-40 for 30 sec. A total of 10 µl of

the MetaSystems DAPI/antifade were applied and the slides were overlayed with a 24 × 60 mm^2 coverslip. The samples were analyzed by a specialist under a fluorescence microscopy. The pictures were captured using imaging software Ikaros (MetaSystems).

^{18}F-FDG PET/CT image acquisition

^{18}F-FDG PET/CT was routinely performed for staging purpose at diagnosis and at one months after therapy as per standard protocol. Serum glucose levels of all patients were confirmed to be less than 150 mg/dL after fasting for at least 6 hours. ^{18}F-FDG PET/CT scans were performed with a 16-row hybrid PET/CT scanner (Discovery IQ Gen 2, GE Healthcare, Milwaukee, Wisconsin, USA). The Body Mass Index (BMI) was calculated for each patient and multiplicated by 0.3 to obtain the total amount (± 10%) of activity in miliCuries (mCi) of ^{18}F-FDG. The dose was injected intravenously 50–80 minutes before PET/CT scanning. All patients were scanned in the supine position with arms elevated above the head. For each patient, an unenhanced CT was performed for anatomic information and attenuation correction (CT scanning parameters: 80 mA, 120 kVp, pixel spacing of 1.367 mm, slice thickness and spacing between slices of 2.5 mm, and a pitch of 1.375). CT images were reconstructed to a 512 × 512 matrix. A three-dimensional (3D) PET scan of the same region was subsequently obtained without any change in position. Emission data were acquired for 130 seconds per bed position, and a total of 4–8 bed positions were performed. PET images were reconstructed in a 192 × 192 matrix and a Bayesian penalized likelihood (BPL) iterative PET reconstruction with a penalization factor or β-value of q350 [53-55].

Image processing and standard imaging biomarkers

All PET/CT images were reviewed by a radiologist and a nuclear medicine physician, both with more than 15 years of experience. The PET/CT images were transferred to the Quibim Precisionv2.8 platform (Quibim SL, Valencia, Spain) for reading performance and lesion segmentation. Considering the advanced Ann Arbor stage of the majority of the patients (**Table 1**), indicating a high spread of the disease, for each patient, all lesions were measured as a whole, thus trying to reflect their tumor burden. The measurement of conventional PET parameters (SUV statistics, MTV, total lesion glycolysis [TLG]) and radiomic features was subsequently performed. Physiological uptakes in organs and tissues like bowel, bladder, brain, injection site were manually removed. In the lesion's volumetry analysis, MTV and TLG were calculated. The SUV_{max} and MTV were automatically generated by the Quibim platform after enclosing each lesion in a cropping sphere. MTV was defined as the volume of voxels with SUVs higher than the threshold of 41% × SUV_{max} [56].

Clinical and genomic features

The following clinical data were collected: sex, age, B symptoms, bulky disease, R-IPI, extra nodal sites, number of extranodal sites, LDH, Ann Arbor stage, human immunodeficiency virus (HIV) status and Easter Cooperative Oncology Group (ECOG). Age, extranodal sites, number of extranodal sites, LDH, Ann Arbor stage, and ECOG performance status were excluded from analyses because they provide redundant information already contained in R-IPI stage. Genomic features included *MYC*/8q24 rearrangement, *MYC*/8q24 amplification, *BCL2*/18q21 rearrangement, *BCL2*/18q21 amplification, *BCL6*/3q27 rearrangement and *BCL6*/3q27 amplification.

Radiomics analysis

The radiomics features were obtained by the Texture Analysis plug-in available in Quibim Precision platform. For the extraction of radiomics features, first order histogram descriptors (skewness, kurtosis, entropy, volume, and max-diameter), as well as second order features were extracted after computing the Gray-Level Co-occurrence Matrix (GLCM), Gray-Level Run Length Matrix (GLRLM), Gray-Level Size Zone Matrix (GLSZM), and the Neighboring Gray-Tone Difference Matrix (NGTDM). Radiomic features calculated by this module comply with the Image Biomarker Standardization Initiative (IBSI) [57].

The Z-score of all imaging features was calculated, and a multivariate analysis was performed after calculating the intra-class correlation coefficients (ICC) to reduce redundant variables.

Table 1. Patient baseline characteristics and response to treatment

Characteristics, *n* (%)	*N* = 33	Median [IQR] time (months)
Diagnosis according to the WHO classification [48]		
GCB DLBCL	13 (39.39)	
Non-GCB	7 (21.21)	
High-grade BCL, NOS	5 (15.15)	
T-cell rich large BCL	3 (9.09)	
Primary mediastinal large BCL	2 (6.06)	
Burkitt lymphoma	2 (6.06)	
High-grade BCL, with MYC and BCL2 and/or BCL6 translocations	1 (2.4)	
B symptoms	22 (66.66)	
LDH > ULN*	22 (66.66)	
ECOG performance status >1	5 (15.15)	
Bulky disease	14 (42.42)	
Ann Arbor stage		
I	3 (9.09)	
II	7 (21.21)	
III	6 (18.18)	
IV	17 (51.52)	
Extranodal involvement > one site	10 (30.3)	
HIV^+	2 (6.06)	
R-IPI risk		
Very good	14 (42.42)	
Good	16 (48.48)	
Poor	3 (9.09)	
Front-line treatment		
Group 1		
R-CHOP or R-CHOP-like regimen	23 (68.43)	
Group 2		
High doses of chemotherapy	10 (29.04)	
Lines of treatment		
1	22 (66.66)	
2	6 (18.18)	
>3	5 (15.15)	
Autologous stem cell transplant	7 (21.21)	
CAR T-cell therapy	2 (6.06)	
Response to treatment		
Long-term complete response	23 (69.69)	72 [52–87]
Late relapse	3 (11.54)**	69 [53–70]
Partial response	2 (6.06)	40 [13–67]
Stable or refractory disease	5 (15.15)	13 [0–82]

*UNL = 480 IU/L. **This percentage is out of the total number of patients who presented complete response at the end of treatment (*n* = 26). CAR-T, Chimeric Antigen Receptor; CBL, B-Cell Lymphoma; ECOG, Easter Cooperative Oncology Group; DLCBL, Diffuse Large B-Cell Lymphoma; GCB, Germinal Center B-Cell-Like; HIV, Human Immunodeficiency Virus; IQR, Interquartile Range; LDH, Lactate Dehydrogenase; NOS, Not Otherwise Specified; R-CHOP, Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone; R-IPI, Revised International Prognostic Index; ULN, Upper Limit of Normal; WHO, World Health Organization.

Outcome and predictors

Multivariate logistic regression predictive models for response to first-line treatment trained with clinical and imaging features (radiomic features + PET conventional parameters), or with clinical, imaging, and genomic features were developed. All clinical and genomic variables

were included as possible predictors unless they provided redundant information as explained, and/or contained missing values and were, therefore, highly unbalanced in the study population. For the selection of imaging features, two different approaches were followed:

• Manual variable selection: the most relevant features according to Mann Whitney Wilcoxon and simple logistic regression test results, and according to visual exploration in univariate analyses were selected.

• Supervised dimensionality reduction through linear discriminant analysis (LDA): a linear classifier generated by fitting the conditional densities of the data, as well as Bayes' rule were used to reduce the input dimensionality towards the most discriminative direction depending on the evaluated response variable (response to first-line treatment).

Statistical analysis

All statistical analyses were performed with Python v.3.8.12, R v.4.2.0 and RStudio. A univariate analysis was conducted for an initial evaluation of the data. This assessment included chi-squared, Fisher and Cramer's V tests for clinical and genomic variables. The correlation of imaging features was analyzed through the Spearman test. Differences between responders and non-responders were evaluated with the Mann Whitney Wilcoxon test and the association of these variables with response to first-line treatment was calculated with a simple logistic regression.

To evaluate predictive model performance different metrics were calculated: area under the curve (AUC), sensitivity, specificity, balanced accuracy, accuracy, and confusion matrix. Models were compared by performing a DeLong test.

A nomogram for prediction of response to first-line treatment (binary outcome) was constructed based upon the multivariate regression model using the statistically significant variables. Thus, the underlying logistic model is given by the equation:

$$Probability\ (response\ to\ first\text{-}line\ treatment) = \frac{e^{(\beta_0+\beta_{pred1}pred1+\beta_{pred2}pred2+\ldots)}}{1+e^{(\beta_0+\beta_{pred1}pred1+\beta_{pred2}pred2+\ldots)}}$$

Beta coefficients were estimated for each covariate and converted to odds ratios as a measure of effect. To obtain the predicted probability of the event (response to first-line treatment), the above equation was calculated using patient's individual characteristics and the beta coefficients derived from the model.

A *P*-value less than 0.05 was considered statistically significant.

Results

Demographic, and clinical characteristics of patients

A total of 33 patients were eligible and included in the study. Of them, 17 (51.52%) were men and 16 women (48.48%), and their median [interquartile range] age at diagnosis was 58 [49–69] years.

Clinical characteristics are summarized in **Table 1**. The majority of patients were diagnosed with germinal center B-cell-like (GCB) DLBCL (*n* = 13, 39.39%) and had B symptoms (*n* = 22, 66.66%). Most patients were in advanced Ann Arbor stage IV (*n* = 17, 51.52%) and 10 (30.3%) had more than one extranodal disease site involvement. According to the R-IPI score, 14 (42.42%) patients were classified as "very good" risk group, 16 (48.48%) as "good" and 3 (9.09%) as "poor". R-CHOP or R-CHOP-like regimens were the most frequently administered front-line therapies (*n* = 21, 63.63%). Most patients only received first-line therapy (*n* = 23, 68.43%).

Median [interquartile range] follow-up of the alive patients was 63, 5 [52–82] months. At the last follow-up, 23 (69.69%) patients achieved long-term complete response (**Table 1**).

Analysis of clinical, imaging and genomic variables

Firstly, categorical variables were analyzed. The distribution of clinical and genomic variables according to the presence or absence of complete response, is shown in Table S1. Notably, only *MYC/8q24 rearrangement* was markedly unbalanced. Indeed, when the association for all categorical variables was measured, that variable resulted highly unstable, and was consequently discarded in subsequent analyses (Figure S1). For the remaining variables, only

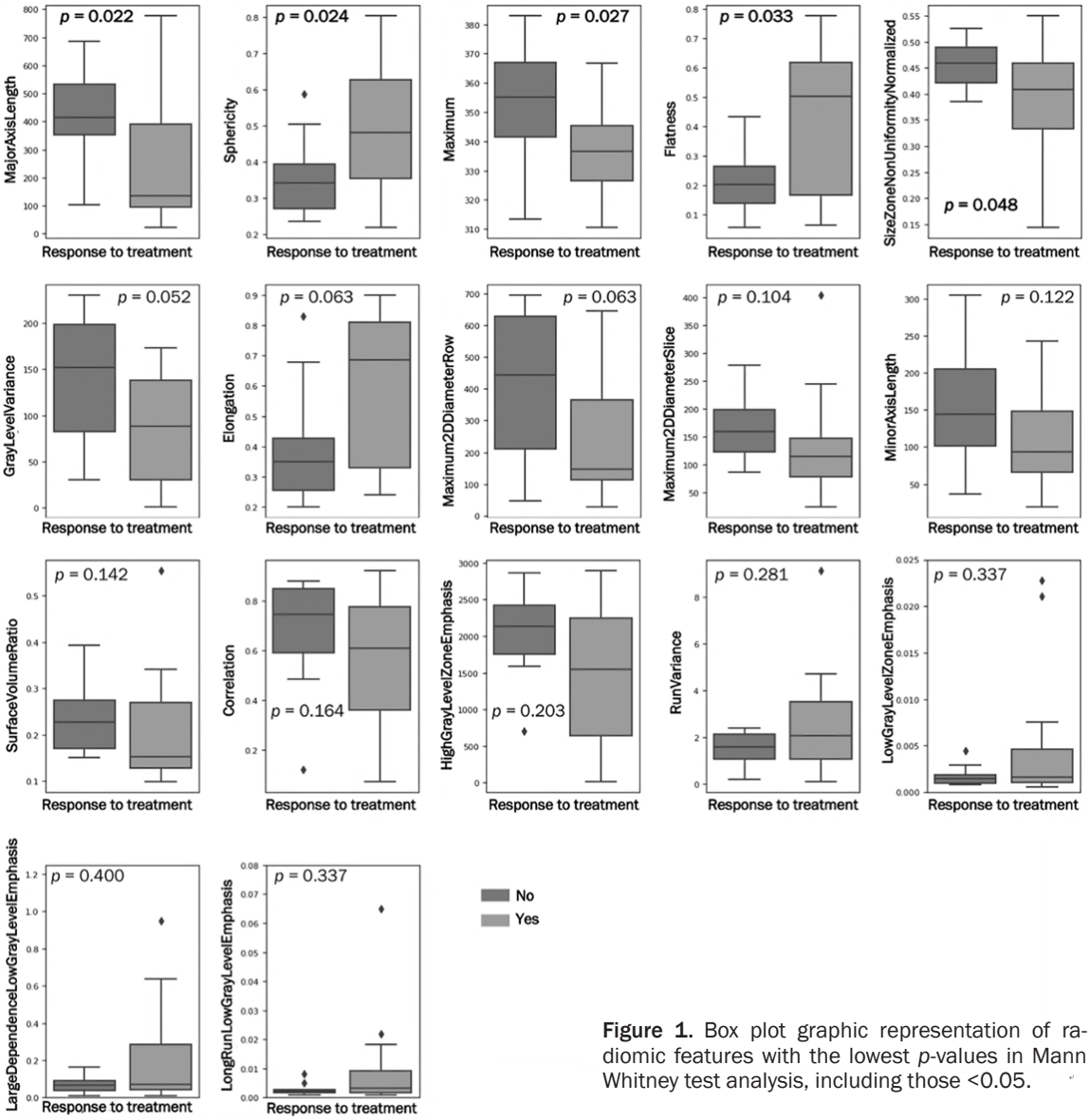

Figure 1. Box plot graphic representation of radiomic features with the lowest *p*-values in Mann Whitney test analysis, including those <0.05.

statistically significant differences were observed in *BCL6/3q27 amplification* (*P* = 0.049) and *BCL2/18q21 amplification* (*P* = 0.05) (Table S1). Notably, these two genomic features showed moderate association with response to treatment (0.25 and 0.45, respectively; Figure S1).

A total of 108 imaging features at diagnosis were analyzed (105 radiomic variables + three PET conventional parameters [SUV_{max}, MTV and TLG]). The most representative ones corresponded to radiomic features and are graphically presented in **Figure 1**. The analysis revealed significant differences especially in variables related with shape and size, such as *sphericity* (*P* = 0.024) and *major axis length* (*P* = 0.033).

Model development

All clinical variables were included in the predictive models. For models including genomic variables, all of them, except for *MYC/8q24 rearrangement* were considered. As detailed in the Methods section, for the selection of imaging variables, an additional simple logistic regression to measure their association with response to treatment was performed. Variables showing remarkable differences between responders

Table 2. Radiomic features finally included in predictive models

Variable	Feature type	Mann Whitney Wilcoxon *P*-value	Logistic regression Pseudo R^2
Elongation	Shape	0.063	0.115
Flatness	Shape	0.033	0.166
MajorAxisLength	Shape	0.022	0.116
Maximum2DDiameterRow	Shape	0.063	0.103
Maximum2DDiameterSlice	Shape	0.104	0.034
MinorAxisLength	Shape	0.122	0.083
Sphericity	Shape	0.024	0.140
SurfaceVolumeRatio	Shape	0.142	0.018
Maximum	First order	0.027	0.171
Correlation	GLCM	0.164	0.040
LongRunLowGrayLevelEmphasis	GLRLM	0.337	0.087
RunVariance	GLRLM	0.281	0.057
GrayLevelVariance	GLSZM	0.052	0.127
HighGrayLevelZoneEmphasis	GLSZM	0.203	0.069
LowGrayLevelZoneEmphasis	GLSZM	0.337	0.075
SizeZoneNonUniformityNormalized	GLSZM	0.048	0.119
LargeDependenceLowGrayLevelEmphasis	GLSZM	0.400	0.085

GLCM, Gray-Level Co-occurrence Matrix; GLDM, Gray-Level Dependence Matrix; GLRLM, Gray-Level Run Length Matrix; GLSZM, Gray-Level Size Zone Matrix.

and non-responders (even not statistically significant) as previously described, as well as those that were able to explain at least 5% response to treatment according to their pseudo R^2 values were finally considered for predictive models. All of them were radiomic features (**Table 2**).

To train the models developed based on the data from the 33 participants included in the study, the following steps were followed:

1) Elimination of multicollinearity.

2) Creation of a first logistic regression to identify possible influential records and subsequently remove them from the model by the residual analysis of the fit.

3) Feature selection by minimizing prediction error (Akaike information criterion [AIC]) with the stepwise method.

4) Selection and evaluation of the final model:

- Hosmer-Lemeshow test to assess the goodness of fit.

- Durbin Watson test to measure residual independence.

- Obtention of an optimal classification threshold.

- Calculation of performance metrics.

Model specification and performance

Predictive model including clinical and imaging features: Table S2 summarizes main performance metrics of the clinical + imaging predictive model built by either applying a manual variable selection process or by applying the LDA reduction method. Overall, metrics were better for the LDA model, which also showed a lower prediction error (AIC = 27.309). As observed in **Table 3**, two radiomic features, *GLSZM_GrayLevelVariance* and *Sphericity*, contributed to explain response to first-line treatment in the manual model ($P = 0.048$, and $P = 0.027$, respectively), while in the LDA model, statistical significance was reached by a variable (named as *LDA*) resulting from the dimensional reduction of all imaging features ($P = 0.034$). Confusion matrices are shown in **Figure 2A**.

Predictive model including clinical, imaging, and genomic features: As observed in Table S3, both manual and LDA models trained with clini-

Table 3. Main characteristics of predictive models trained with clinical and radiomic variables and built with manual variable selection or by applying the linear discriminant analysis reduction method

Variable	Predictive model (Manual selection)			Predictive model (LDA)		
	OR	95% CI	*P*-value	OR	95% CI	*P*-value
Intercept	2.29	0.51–13.2	0.074	3.45	0.87–26.58	0.136
B symptoms	13.33	1.08–507.76	0.085	-	-	-
GLSZM_GrayLevelVariance	0.23	0.037–1.26	**0.048**	-	-	-
Maximum2DDiameterSlice	4.61	0.93–52.98	0.121	-	-	-
Sphericity	14.74	2.16–301.87	**0.027**	-	-	-
Bulky	-	-	-	6.52	0.65–164.02	0.151
LDA	-	-	-	10.69	2.23–223.63	**0.034**
Validation						
Hosmer-Lemeshow test (*P*-value)	-	-	0.839	-	-	0.957
Durbin Watson test (*P*-value)	-	-	0.838	-	-	0.802
Classification threshold	-	-	0.529	-	-	0.426

CI, Confidence Interval; GLSMZ: Gray-Level Size Zone Matrix; LDA, Linear Discriminant Analysis; OR, Odds Ratio. Bold statistically significant results.

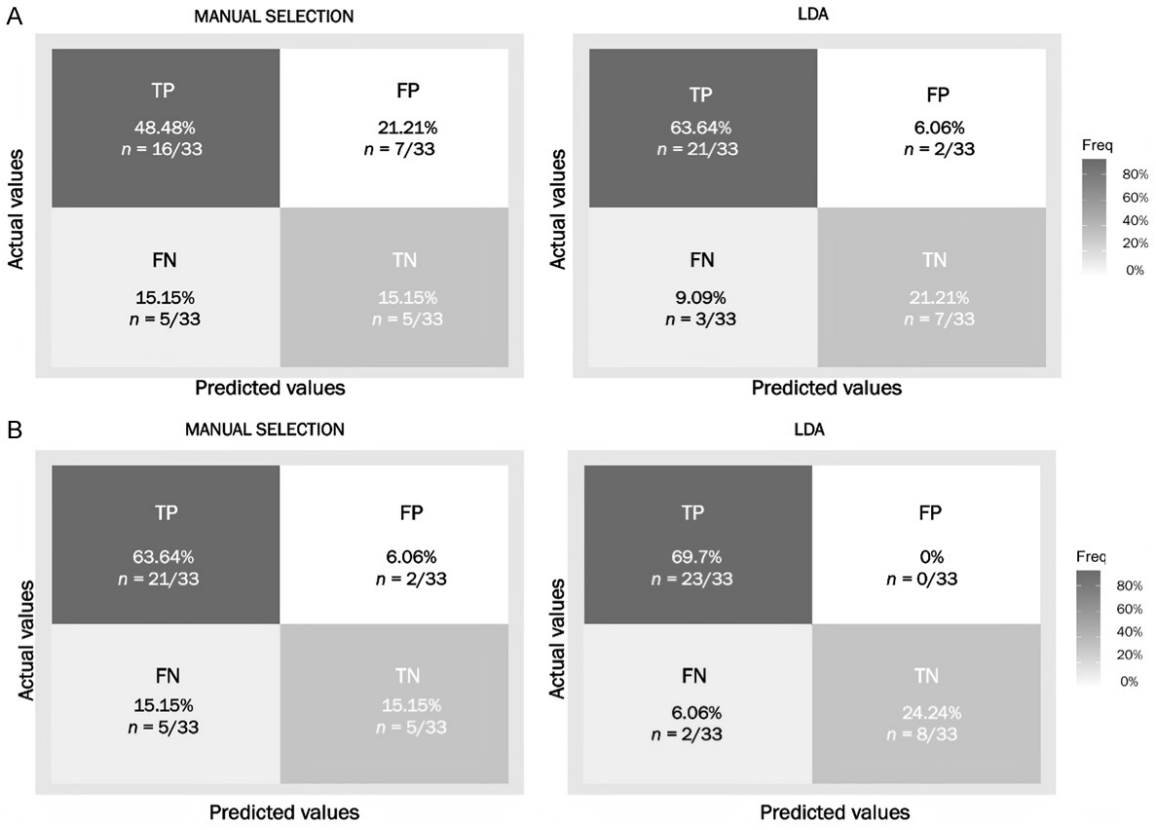

Figure 2. Confusion matrices for manual selection and LDA predictive models trained with clinical and radiomic variables (A) or with clinical, radiomic and genomic variables (B). FN, False Negative; FP, False Positive; LDA, Linear Discriminant Analysis; TN, True Negative; TP, True Positive.

cal, imaging, and genomic features were able to predict response with a balanced accuracy superior to 80%, although metrics for the LDA model were generally better, achieving a 90% balanced accuracy, a sensitivity of 100% and a specificity of 80% with a lower prediction error (AIC = 27.451). **Table 4** summarizes the main characteristics of both models. In both of them,

Table 4. Main characteristics of predictive models trained with clinical, radiomic and genomic variables and built with manual variable selection or by applying the linear discriminant analysis reduction method

Variable	Predictive model (Manual selection)			Predictive model (LDA)		
	OR	95% CI	*P*-value	OR	95% CI	*P*-value
Intercept	5.75	1.92–23.10	**0.005**	12.02	2.61–121.51	**0.008**
BCL6 amplification	0.06	0.0033–0.46	**0.018**	0.08	0.0042–0.73	**0.045**
GLCM_Correlation	0.34	0.094–0.88	**0.05**	-	-	-
LDA	-	-	-	0.21	-0.046–0.53	**0.008**
Validation						
Hosmer-Lemeshow test (*P*-value)	-	-	0.773	-	-	0.253
Durbin Watson test (*P*-value)	-	-	0.244	-	-	0.150
Classification threshold	-	-	0.424	-	-	0.424

CI, Confidence Interval; GLCM: Gray-Level Co-occurrence Matrix; LDA, Linear Discriminant Analysis; OR, Odds Ratio. Bold statistically significant results.

Table 5. Summary of performance metrics for all the predictive models and comparisons between models

Metrics	Manual selection		LDA		*P*-value*
	Model 1 (clinical + radiomic features)	Model 2 (clinical + radiomic + genomic features)	Model 1 (clinical + radiomic features)	Model 2 (clinical + radiomic + genomic features)	
AIC	35.697	27.309	34.816	27.451	
AUC	0.598	0.757	0.891	0.904	
Sensitivity	69.6%	95.2%	91.3%	100.0%	
Specificity	50.0%	50%	70.0%	80.0%	
Balanced accuracy	59.8%	72.6%	80.7%	90.0%	
Accuracy	63.6%	80.6%	84.9%	93.9%	
Comparisons					
Manual selection vs LDA (model 1)					0.6344
Manual selection vs LDA (model 2)					0.2242
Model 1 vs Model 2 (manual selection)					0.2815
Model 1 vs Model 2 (LDA)					0.8305

*DeLong test of ROC curve of models. AIC, Akaike Information Criterion; AUC, Area Under the Curve; LDA, Linear Discriminant Analysis.

BCL6 amplification significantly contributed to explain response to first-line treatment, indicating that patients harboring that mutation are less likely to respond. Additionally, imaging variables also retained prediction ability, with a statistically significant contribution of the radiomic *GLCM_Correlation* variable in the manual model ($P = 0.05$), and of the *LDA* variable in the LDA model; a variable resulting from the dimensional reduction of all imaging features ($P = 0.008$). Overall, classification was good with both models (**Figure 2B**).

Table 5 summarizes the performance metrics for all the models developed, as well as the comparisons between them. Overall, for both the models trained without genomic features and for those including them, the LDA dimensionality reduction approach led to better metrics, with higher AUC and accuracy values, although no statistical differences were found between them. Importantly, the inclusion of genomic features clearly improved the predictive ability of the models, being the LDA model trained with clinical, imaging and genomic features, the one yielding the highest performance metrics, with an AUC of 0.904, a balanced accuracy of 90%, a sensitivity of 100% and a specificity of 80%. A nomogram utilizing the *BCL6 amplification* and *LDA* variable results in the prediction of treatment response was created, with *BCL6 amplification* and *LDA* variable as additive cofactors and their contribution to the prediction based upon the coefficients of

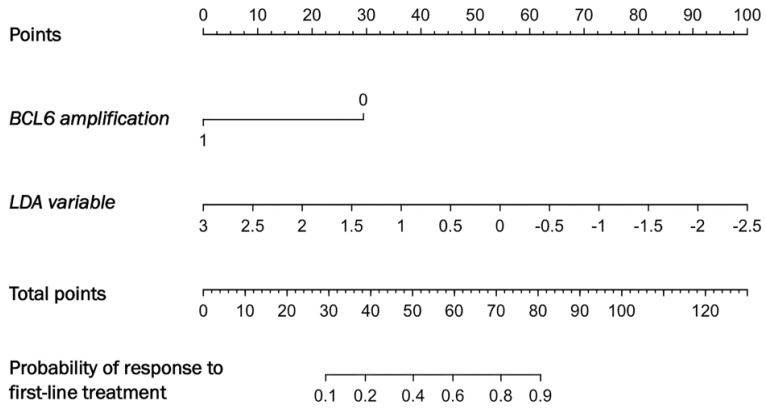

Figure 3. Nomogram for response to first-line treatment in diffuse large B-cell lymphoma (DLBCL) based on *BCL6 amplification* and *LDA variable* results provided by the model combining imaging, clinical and genomic features and developed following the LDA approach. As an example, a patient with no amplification in *BCL6* will receive a score of 30; if their *LDA variable* result is 0.5, this will produce a score of approximately 45, resulting in a total score of 75. On the lower 2 panels, a total of 75 points is estimated to give rise to around 85% probability of response to first-line treatment. Note on the nomogram, *BCL6 amplification* 1 corresponds to presence of this mutation and *BCL6 amplification* 0 corresponds to amplification of this gene not detected.

the logistic regression model developed following the LDA approach (**Figure 3**).

Discussion

After treatment onset, 10%–15% of DLBCL patients have primary refractory disease within 3 months, and early relapse occurs in a further 20%–35% of them [58], even if they achieved complete response after first-line immunochemotherapy [59]. It is therefore necessary to early identify patients who are unlikely to be cured with the standard front-line therapy, so that they could benefit from alternative therapeutic approaches. Several clinical risk indexes built with clinical and biological variables have been proposed to identify those refractory patients, but neither of them fully the identification of this high-risk group. In this study, we demonstrated that a combined model including clinical, imaging, and genomic features was able to successfully predict response to first-line treatment, retaining a higher predictive ability than a model trained only with clinical and imaging variables.

Gene rearrangements involving *MYC*, *BCL2* and/or *BCL6* are common in DLBCL patients, observed in 8%–14%, about 20% and up to 30% of the cases, respectively [48]. Importantly, they have been identified as poor prognostic factors in the disease, especially for *MYC* when occurring in combination with *BCL2* and/or *BCL6* [15-17, 60-62]. In a similar way, although still controversial, some studies suggest that the amplification of at least *MYC* and *BCL2*, could also predict poor prognosis in DLBCL [21, 22, 63]. Interestingly, in our series, and regardless of the approach followed to build the model, we have identified the amplification of *BCL6* gene as one of the main features helping to predict response to first-line treatment. To date, the available evidence about the impact of this genomic alteration on patient's outcomes is scarce [19, 63, 64]; only one study by Willenbacher et al. [64] has demonstrated its predictive value in terms of survival and disease progression. Thus, when treated with standard chemotherapy, patients with copy number gains had a significantly worse overall survival (OS) and recurrence free survival (RFS) compared with all other patients (double or triple hit DLBCL or patients with no significant alterations). In agreement with this study, our results highlight that the presence of a *BCL6* amplification is associated with an absence of response to front-line therapy in DLBCL patients. The robustness of this result is supported, not only by its statistical significance, but also by its relevance when the patient population is analyzed in detail. Thus, in our cohort less than a half of the patients (n = 11/33, 33%) had *BCL6* amplification, of whom only five (15.15%) responded to treatment. Among those five patients, in two of them *BCL6 amplification* was the only genomic alteration detected. Overall, this stresses the potent contribution of this amplification to the predictive model despite the relative low presence of patients harboring this mutation in the study population. This is also in agreement with the study by Willenbacher et al. [64], in which patients with copy number gains accounted for 10% of the total, 25% of which presented BCL6 amplification.

Radiomics is a quantitative approach experiencing a rapid growth in the last decade, and whose potential in diagnosis and treatment outcome prediction in oncology is unquestionable [65]. This is supported by several studies in which radiomic-based models allow prediction of different clinical outcomes (e.g., OS, progression-free survival [PFS], and recurrence) in various cancer types [65]. Specifically in DLBCL, several studies have demonstrated the predictive value of radiomics. Thus, a preliminary work carried out by our group [66] already evidenced how the addition of radiomics features to a predictive model based on the conventional IPI evaluation of patients, significantly increased the performance to identify patients requiring more than one treatment lines or likely to respond to treatment. Other authors have also assessed the performance of radiomics for predicting response, evidencing that tumor textural heterogeneity is associated with treatment failure [47, 67]. Here, we have confirmed the predictive value of radiomics through conventional pre-treatment ^{18}F-FDG PET/CT scans; the standard of care imaging modality for DLBCL patients [24-27]. Regardless of the approach followed, among imaging features, only radiomic variables significantly contributed to explain treatment response even more than clinical variables, which were not even included in the models as important predictive factors when genomic characteristics were considered. Additionally, conventional PET parameters did not retain any predictive value, and no differences were found between responders and non-responders in the exploratory analysis either. When applying the manual variable selection method, shape- and heterogeneity-related radiomic features, *GLSZM_GrayLevelVariance* and *Sphericity* (in the clinical + imaging model) and *GLCM_Correlation* (in the clinical + imaging + genomic model), resulted statistically significant, in line with previous reports. However, according to the lesion segmentation and assessment process followed in this study, reflecting tumor burden instead of individual lesions, these results would be indicative of lesion dissemination rather than of tumor heterogeneity. Thus, lower values of *GLSZM_GrayLevelVariance* and a higher *sphericity* in the group of responders, as observed in our cohort, would suggest both a lower lesion distribution and a lower metabolic heterogeneity within lesions, with a more homogeneous radiotracer uptake, which would consequently favor response to treatment. Indeed, the distance between the two lesions that are the furthest apart (Dmax), a simple imaging feature measured from FDG PET scans that reflects lesions dissemination has been recently introduced in DLBCL patients [44]. Hence, large lesion dissemination measured with this parameter has been found to be a strong prognostic factor for shorter PFS and OS in DLBCL patients, stressing the importance of capturing lymphoma extension. Here, we provide evidence on radiomic features as potential imaging biomarkers to measure lesion dissemination, and subsequently to predict treatment response. In addition, our results demonstrate that the potential of radiomics is far beyond the predictive value of some specific features. As observed, by applying a dimensionality reduction approach, we found that the resulting variable (referred as *LDA*), reflecting the whole set of imaging variables analyzed, mainly composed of radiomic features, significantly contributed to explain treatment response, both in the model excluding and including genomic features, and in both cases showing better performance metrics than the manual models. This strongly suggests that a panel of imaging features and especially of radiomic features, could potentially serve to guide therapeutic strategy in the future.

Importantly, our study stresses the relevance of genomic features when predicting treatment response. Thus, the highest AUC, balanced accuracy, sensitivity, and specificity (0.904, 90%, 100% and 80%, respectively) was achieved when these variables were combined together with clinical and imaging features in a LDA model. We believe that the lack of statistical significance when model comparisons were performed may be due to the limited size of our study population. However, it is important to note that these comparisons only take AUC values into account, and that overall, performance metrics, as explained, were superior for LDA-based models. To date, only one congress abstract published in 2021 has assessed the utility of ^{18}F-FDG PET/CT radiomics features in combination with genomic data in identifying DLBCL patients at high risk for relapse [68]. The authors concluded that the positive predictive value increased when radiomics features were added to the clinical and genetic parame-

ters. The optimal combined model included *MYC* status, WHO performance status, LDH, and different conventional PET features including Dmax. It is worth mentioning that PET conventional parameters and dissemination variables were considered as radiomic features in this study, not matching the concept as defined by Pyradiomics website [69] and widely used in the literature. Additionally, gene amplifications were not quantified. Although methodological differences impede us to make comparisons, the results of this study clearly strengthen the notion of combined models as powerful tools to predict treatment response, again highlighting the importance of lesion dissemination.

Finally, we were able to construct a nomogram that, by the combination of *BCL6 amplification* and *LDA* variable as additive cofactors, was able to successfully predict the probability of response to first-line treatment. Despite its good results and promising application, further testing along with the general model validation is required and will be performed in future studies.

However, this study has also some limitations. Firstly, it was a single-center study with limited number of patients included and its application to data/patients from larger series in other institutions should be further explored. Secondly, given the unbalance of the dataset in terms of *MYC* rearrangements, it was impossible to include this relevant genomic feature in DLBCL in the analysis. Thirdly, regarding predictive models, it is important to note that all metrics provided correspond to training metrics, as the adjusted models were developed including all available data given the limited cohort size. Consequently, our results must be interpreted cautiously; the predictive ability of our models with new datasets will need to be further validated. Finally, we are aware that the presence of different treatments and disease subtypes may be affecting the outcomes. Regrettably, because of the small sample size, we were unable to run independent analyses for each of the different subtypes. We will address all these limitations in future studies.

In conclusion, our study demonstrates that a combination of clinical, imaging and genomic features enables to successfully predict response to first-line therapy in DLBCL patients, with the amplification of *BCL6* as the genetic marker retaining the highest predictive value. Additionally, a panel of imaging features, and in particular of radiomic data, extracted from a routinely PET/CT scan at baseline may provide important information when predicting treatment response, with lesion dissemination-related radiomic features deserving especial attention. This consideration might be particularly interesting for those hospitals and health centers where costs may limit genetic assessments. In summary, the combination of molecular and imaging characteristics at diagnosis could lead to a more accurate selection of patients, to increase tailor therapy. A nomogram could be helpful in the identification of high-risk patients for new therapeutic approaches.

Acknowledgements

María José Terol received funding for this work by the Gilead Fellowship project number GLD21/00088. Ángel Alberich-Bayarri is the CEO and a shareholder of Quibim S.L. Irene Mayorga-Ruiz is a shareholder of Quibim S.L.

Disclosure of conflict of interest

None.

Address correspondence to: María José Terol, Hematology Department, Hospital Clínico Universitario-INCLIVA, Avenue Blasco Ibáñez 17, Valencia 46010, Spain. Tel: +34-961-97-35-00; +34-627-91-01-98; E-mail: maria.jose.terol@uv.es

References

[1] Liu Y and Barta SK. Diffuse large B-cell lymphoma: 2019 update on diagnosis, risk stratification, and treatment. Am J Hematol 2019; 94: 604-616.

[2] Thandra KC, Barsouk A, Saginala K, Padala SA, Barsouk A and Rawla P. Epidemiology of non-Hodgkin's lymphoma. Med Sci (Basel) 2021; 9: 5.

[3] Friedberg JW and Fisher RI. Diffuse large B-cell lymphoma. Hematol Oncol Clin North Am 2008; 22: 941-952, ix.

[4] Coiffier B and Sarkozy C. Diffuse large B-cell lymphoma: R-CHOP failure-what to do? Hematology Am Soc Hematol Educ Program 2016; 2016: 366-378.

[5] Coiffier B, Lepage E, Briere J, Herbrecht R, Tilly H, Bouabdallah R, Morel P, Van Den Neste E, Salles G, Gaulard P, Reyes F, Lederlin P and Gisselbrecht C. CHOP chemotherapy plus ritux-

imab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. N Engl J Med 2002; 346: 235-242.
[6] Friedberg JW. Relapsed/refractory diffuse large B-cell lymphoma. Hematology Am Soc Hematol Educ Program 2011; 2011: 498-505.
[7] Tilly H, Gomes da Silva M, Vitolo U, Jack A, Meignan M, Lopez-Guillermo A, Walewski J, André M, Johnson PW, Pfreundschuh M and Ladetto M; ESMO Guidelines Committee. Diffuse large B-cell lymphoma (DLBCL): ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2015; 26 Suppl 5: v116-125.
[8] Bari A, Marcheselli L, Sacchi S, Marcheselli R, Pozzi S, Ferri P, Balleari E, Musto P, Neri S, Aloe Spiriti MA and Cox MC. Prognostic models for diffuse large B-cell lymphoma in the rituximab era: a never-ending story. Ann Oncol 2010; 21: 1486-1491.
[9] Ruppert AS, Dixon JG, Salles G, Wall A, Cunningham D, Poeschel V, Haioun C, Tilly H, Ghesquieres H, Ziepert M, Flament J, Flowers C, Shi Q and Schmitz N. International prognostic indices in diffuse large B-cell lymphoma: a comparison of IPI, R-IPI, and NCCN-IPI. Blood 2020; 135: 2041-2048.
[10] Sehn LH, Berry B, Chhanabhai M, Fitzgerald C, Gill K, Hoskins P, Klasa R, Savage KJ, Shenkier T, Sutherland J, Gascoyne RD and Connors JM. The revised international prognostic index (R-IPI) is a better predictor of outcome than the standard IPI for patients with diffuse large B-cell lymphoma treated with R-CHOP. Blood 2007; 109: 1857-1861.
[11] Tay K, Tai D, Tao M, Quek R, Ha TC and Lim ST. Relevance of the international prognostic index in the rituximab era. J Clin Oncol 2011; 29: e14; author reply e15.
[12] Zhou Z, Sehn LH, Rademaker AW, Gordon LI, Lacasce AS, Crosby-Thompson A, Vanderplas A, Zelenetz AD, Abel GA, Rodriguez MA, Nademanee A, Kaminski MS, Czuczman MS, Millenson M, Niland J, Gascoyne RD, Connors JM, Friedberg JW and Winter JN. An enhanced international prognostic index (NCCN-IPI) for patients with diffuse large B-cell lymphoma treated in the rituximab era. Blood 2014; 123: 837-842.
[13] Riedell PA and Smith SM. Double hit and double expressors in lymphoma: definition and treatment. Cancer 2018; 124: 4622-4632.
[14] Sesques P and Johnson NA. Approach to the diagnosis and treatment of high-grade B-cell lymphomas with MYC and BCL2 and/or BCL6 rearrangements. Blood 2017; 129: 280-288.
[15] Akyurek N, Uner A, Benekli M and Barista I. Prognostic significance of MYC, BCL2, and BCL6 rearrangements in patients with diffuse large B-cell lymphoma treated with cyclophosphamide, doxorubicin, vincristine, and prednisone plus rituximab. Cancer 2012; 118: 4173-4183.
[16] Barrans S, Crouch S, Smith A, Turner K, Owen R, Patmore R, Roman E and Jack A. Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab. J Clin Oncol 2010; 28: 3360-3365.
[17] Johnson NA, Slack GW, Savage KJ, Connors JM, Ben-Neriah S, Rogic S, Scott DW, Tan KL, Steidl C, Sehn LH, Chan WC, Iqbal J, Meyer PN, Lenz G, Wright G, Rimsza LM, Valentino C, Brunhoeber P, Grogan TM, Braziel RM, Cook JR, Tubbs RR, Weisenburger DD, Campo E, Rosenwald A, Ott G, Delabie J, Holcroft C, Jaffe ES, Staudt LM and Gascoyne RD. Concurrent expression of MYC and BCL2 in diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. J Clin Oncol 2012; 30: 3452-3459.
[18] Rosenwald A, Bens S, Advani R, Barrans S, Copie-Bergman C, Elsensohn MH, Natkunam Y, Calaminici M, Sander B, Baia M, Smith A, Painter D, Pham L, Zhao S, Ziepert M, Jordanova ES, Molina TJ, Kersten MJ, Kimby E, Klapper W, Raemaekers J, Schmitz N, Jardin F, Stevens WBC, Hoster E, Hagenbeek A, Gribben JG, Siebert R, Gascoyne RD, Scott DW, Gaulard P, Salles G, Burton C, de Jong D, Sehn LH and Maucort-Boulch D. Prognostic significance of MYC rearrangement and translocation partner in diffuse large B-cell lymphoma: a study by the lunenburg lymphoma biomarker consortium. J Clin Oncol 2019; 37: 3359-3368.
[19] Quesada AE, Medeiros LJ, Desai PA, Lin P, Westin JR, Hawsawi HM, Wei P, Tang G, Seegmiller AC, Reddy NM, Yin CC, Wang W, Xu J, Miranda RN, Zuo Z and Li S. Increased MYC copy number is an independent prognostic factor in patients with diffuse large B-cell lymphoma. Mod Pathol 2017; 30: 1688-1697.
[20] Jiang HY, Li HL, Hu H, He Y and Zhao T. Detection of t (14; 18) translocation and bcl-2 amplification in diffuse large B-cell lymphoma. Zhonghua Bing Li Xue Za Zhi 2007; 36: 84-89.
[21] Yoon SO, Jeon YK, Paik JH, Kim WY, Kim YA, Kim JE and Kim CW. MYC translocation and an increased copy number predict poor prognosis in adult diffuse large B-cell lymphoma (DLBCL), especially in germinal centre-like B cell (GCB) type. Histopathology 2008; 53: 205-217.
[22] Krull JE, Wenzl K, Hartert KT, Manske MK, Sarangi V, Maurer MJ, Larson MC, Nowakowski GS, Ansell SM, McPhail E, Habermann TM, Link BK, King RL, Cerhan JR and Novak AJ.

Somatic copy number gains in MYC, BCL2, and BCL6 identifies a subset of aggressive alternative-DH/TH DLBCL patients. Blood Cancer J 2020; 10: 117.

[23] Lu TX, Fan L, Wang L, Wu JZ, Miao KR, Liang JH, Gong QX, Wang Z, Young KH, Xu W, Zhang ZH and Li JY. MYC or BCL2 copy number aberration is a strong predictor of outcome in patients with diffuse large B-cell lymphoma. Oncotarget 2015; 6: 18374-18388.

[24] Barrington SF, Mikhaeel NG, Kostakoglu L, Meignan M, Hutchings M, Müeller SP, Schwartz LH, Zucca E, Fisher RI, Trotman J, Hoekstra OS, Hicks RJ, O'Doherty MJ, Hustinx R, Biggi A and Cheson BD. Role of imaging in the staging and response assessment of lymphoma: consensus of the international conference on malignant lymphomas imaging working group. J Clin Oncol 2014; 32: 3048-3058.

[25] Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E and Lister TA; Alliance, Australasian Leukaemia and Lymphoma Group; Eastern Cooperative Oncology Group; European Mantle Cell Lymphoma Consortium; Italian Lymphoma Foundation; European Organisation for Research; Treatment of Cancer/Dutch Hemato-Oncology Group; Grupo Español de Médula Ósea; German High-Grade Lymphoma Study Group; German Hodgkin's Study Group; Japanese Lymphorra Study Group; Lymphoma Study Association; NCIC Clinical Trials Group; Nordic Lymphoma Study Group; Southwest Oncology Group; United Kingdom National Cancer Research Institute. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. J Clin Oncol 2014; 32: 3059-3068.

[26] Gómez León N, Delgado-Bolton RC, Del Campo Del Val L, Cabezas B, Arranz R, García M, Cannata J, González Ortega S, Pérez Sáez MÁ, López-Botet B, Rodríguez-Vigil B, Mateo M, Colletti PM, Rubello D and Carreras JL. Multicenter comparison of contrast-enhanced FDG PET/CT and 64-slice multi-detector-row CT for initial staging and response evaluation at the end of treatment in patients with lymphoma. Clin Nucl Med 2017; 42: 595-602.

[27] Voltin CA, Mettler J, Grosse J, Dietlein M, Baues C, Schmitz C, Borchmann P, Kobe C and Hellwig D. FDG-PET imaging for Hodgkin and diffuse large B-cell lymphoma-an updated overview. Cancers (Basel) 2020; 12: 601.

[28] Frood R, Burton C, Tsoumpas C, Frangi AF, Gleeson F, Patel C and Scarsbrook A. Baseline PET/CT imaging parameters for prediction of treatment outcome in Hodgkin and diffuse large B cell lymphoma: a systematic review. Eur J Nucl Med Mol Imaging 2021; 48: 3198-3220.

[29] Adams HJ, de Klerk JM, Fijnheer R, Heggelman BG, Dubois SV, Nievelstein RA and Kwee TC. Prognostic superiority of the national comprehensive cancer network international prognostic index over pretreatment whole-body volumetric-metabolic FDG-PET/CT metrics in diffuse large B-cell lymphoma. Eur J Haematol 2015; 94: 532-539.

[30] Guo B, Tan X, Ke Q and Cen H. Prognostic value of baseline metabolic tumor volume and total lesion glycolysis in patients with lymphoma: a meta-analysis. PLoS One 2019; 14: e0210224.

[31] Mikhaeel NG, Smith D, Dunn JT, Phillips M, Møller H, Fields PA, Wrench D and Barrington SF. Combination of baseline metabolic tumour volume and early response on PET/CT improves progression-free survival prediction in DLBCL. Eur J Nucl Med Mol Imaging 2016; 43: 1209-1219.

[32] Schmitz C, Hüttmann A, Müller SP, Hanoun M, Boellaard R, Brinkmann M, Jöckel KH, Dührsen U and Rekowski J. Dynamic risk assessment based on positron emission tomography scanning in diffuse large B-cell lymphoma: post-hoc analysis from the PETAL trial. Eur J Cancer 2020; 124: 25-36.

[33] Toledano MN, Desbordes P, Banjar A, Gardin I, Vera P, Ruminy P, Jardin F, Tilly H and Becker S. Combination of baseline FDG PET/CT total metabolic tumour volume and gene expression profile have a robust predictive value in patients with diffuse large B-cell lymphoma. Eur J Nucl Med Mol Imaging 2018; 45: 680-688.

[34] Vercellino L, Cottereau AS, Casasnovas O, Tilly H, Feugier P, Chartier L, Fruchart C, Roulin L, Oberic L, Pica GM, Ribrag V, Abraham J, Simon M, Gonzalez H, Bouabdallah R, Fitoussi O, Sebban C, López-Guillermo A, Sanhes L, Morschhauser F, Trotman J, Corront B, Choufi B, Snauwaert S, Godmer P, Briere J, Salles G, Gaulard P, Meignan M and Thieblemont C. High total metabolic tumor volume at baseline predicts survival independent of response to therapy. Blood 2020; 135: 1396-1405.

[35] Adams MC, Turkington TG, Wilson JM and Wong TZ. A systematic review of the factors affecting accuracy of SUV measurements. AJR Am J Roentgenol 2010; 195: 310-320.

[36] Osman MM, Muzaffar R, Altinyay ME and Teymouri C. FDG dose extravasations in PET/CT: frequency and impact on SUV measurements. Front Oncol 2011; 1: 41.

[37] Martín-Saladich Q, Reynés-Llompart G, Sabaté-Llobera A, Palomar-Muñoz A, Domingo-Domènech E and Cortés-Romera M. Comparison of different automatic methods for the delineation of the total metabolic tumor volume in I-II stage Hodgkin lymphoma. Sci Rep 2020; 10: 12590.

[38] Araf S, Korfi K, Bewicke-Copley F, Wang J, Cogliatti S, Kumar E, Forrer F, Barrington SF, Graham TA, Scott DW, Rimsza LM, Davies A, Johnson P, Okosun J, Fitzgibbon J and Fehr M. Genetic heterogeneity highlighted by differential FDG-PET response in diffuse large B-cell lymphoma. Haematologica 2020; 105: 318-321.

[39] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A and Aerts HJ. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012; 48: 441-446.

[40] Stanta G and Bonin S. Overview on clinical relevance of intra-tumor heterogeneity. Front Med (Lausanne) 2018; 5: 85.

[41] Burrell RA, McGranahan N, Bartek J and Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. Nature 2013; 501: 338-345.

[42] McGranahan N and Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. Cell 2017; 168: 613-628.

[43] Aide N, Fruchart C, Nganoa C, Gac AC and Lasnon C. Baseline (18)F-FDG PET radiomic features as predictors of 2-year event-free survival in diffuse large B cell lymphomas treated with immunochemotherapy. Eur Radiol 2020; 30: 4623-4632.

[44] Cottereau AS, Nioche C, Dirand AS, Clerc J, Morschhauser F, Casasnovas O, Meignan M and Buvat I. ^{18}F-FDG PET dissemination features in diffuse large B-cell lymphoma are predictive of outcome. J Nucl Med 2020; 61: 40-45.

[45] Eertink JJ, van de Brug T, Wiegers SE, Zwezerijnen GJC, Pfaehler EAG, Lugtenburg PJ, van der Holt B, de Vet HCW, Hoekstra OS, Boellaard R and Zijlstra JM. (18)F-FDG PET baseline radiomics features improve the prediction of treatment outcome in diffuse large B-cell lymphoma. Eur J Nucl Med Mol Imaging 2022; 49: 932-942.

[46] Frood R, Clark M, Burton C, Tsoumpas C, Frangi AF, Gleeson F, Patel C and Scarsbrook AF. Discovery of pre-treatment FDG PET/CT-derived radiomics-based models for predicting outcome in diffuse large B-cell lymphoma. Cancers (Basel) 2022; 14: 1711.

[47] Santiago R, Ortiz Jimenez J, Forghani R, Muthukrishnan N, Del Corpo O, Karthigesu S, Haider MY, Reinhold C and Assouline S. CT-based radiomics model with machine learning for predicting primary treatment failure in diffuse large B-cell lymphoma. Transl Oncol 2021; 14: 101188-101188.

[48] Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, Advani R, Ghielmini M, Salles GA, Zelenetz AD and Jaffe ES. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood 2016; 127: 2375-2390.

[49] González-Barca E, Coronado M, Martín A, Montalbán C, Montes-Moreno S, Panizo C, Rodríguez G, Sancho JM and López-Hernández A. Spanish lymphoma group (GELTAMO) guidelines for the diagnosis, staging, treatment, and follow-up of diffuse large B-cell lymphoma. Oncotarget 2018; 9: 32383-32399.

[50] Zelenetz AD, Gordon LI, Chang JE, Christian B, Abramson JS, Advani RH, Bartlett NL, Budde LE, Caimi PF, De Vos S, Dholaria B, Fakhri B, Fayad LE, Glenn MJ, Habermann TM, Hernandez-Ilizaliturri F, Hsi E, Hu B, Kaminski MS, Kelsey CR, Khan N, Krivacic S, LaCasce AS, Lim M, Narkhede M, Rabinovitch R, Ramakrishnan P, Reid E, Roberts KB, Saeed H, Smith SD, Svoboda J, Swinnen LJ, Tuscano J, Vose JM, Dwyer MA and Sundar H. NCCN guidelines® insights: B-cell lymphomas, version 5.2021. J Natl Compr Canc Netw 2021; 19: 1218-1230.

[51] Barrington SF, Qian W, Somer EJ, Franceschetto A, Bagni B, Brun E, Almquist H, Loft A, Højgaard L, Federico M, Gallamini A, Smith P, Johnson P, Radford J and O'Doherty MJ. Concordance between four European centres of PET reporting criteria designed for use in multicentre trials in Hodgkin lymphoma. Eur J Nucl Med Mol Imaging 2010; 37: 1824-1833.

[52] Meignan M, Gallamini A, Meignan M, Gallamini A and Haioun C. Report on the first international workshop on interim-PET-scan in lymphoma. Leuk Lymphoma 2009; 50: 1257-1260.

[53] Jaskowiak CJ, Bianco JA, Perlman SB and Fine JP. Influence of reconstruction iterations on 18F-FDG PET/CT standardized uptake values. J Nucl Med 2005; 46: 424-428.

[54] Messerli M, Stolzmann P, Egger-Sigg M, Trinckauf J, D'Aguanno S, Burger IA, von Schulthess GK, Kaufmann PA and Huellner MW. Impact of a Bayesian penalized likelihood reconstruction algorithm on image quality in novel digital PET/CT: clinical implications for the assessment of lung tumors. EJNMMI Phys 2018; 5: 27.

[55] Teoh EJ, McGowan DR, Macpherson RE, Bradley KM and Gleeson FV. Phantom and clinical evaluation of the bayesian penalized likelihood reconstruction algorithm Q.Clear on an LYSO PET/CT system. J Nucl Med 2015; 56: 1447-1452.

[56] Boellaard R, O'Doherty MJ, Weber WA, Mottaghy FM, Lonsdale MN, Stroobants SG, Oyen WJ, Kotzerke J, Hoekstra OS, Pruim J, Marsden PK, Tatsch K, Hoekstra CJ, Visser EP, Arends B, Verzijlbergen FJ, Zijlstra JM, Comans EF, Lammertsma AA, Paans AM, Willemsen AT, Beyer T, Bockisch A, Schaefer-Prokop C, Delbeke D, Baum RP, Chiti A and Krause BJ. FDG PET and

PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. Eur J Nucl Med Mol Imaging 2010; 37: 181-200.

[57] Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, Ashrafinia S, Bakas S, Beukinga RJ, Boellaard R, Bogowicz M, Boldrini L, Buvat I, Cook GJR, Davatzikos C, Depeursinge A, Desseroit MC, Dinapoli N, Dinh CV, Echegaray S, El Naqa I, Fedorov AY, Gatta R, Gillies RJ, Goh V, Götz M, Guckenberger M, Ha SM, Hatt M, Isensee F, Lambin P, Leger S, Leijenaar RTH, Lenkowicz J, Lippert F, Losnegård A, Maier-Hein KH, Morin O, Müller H, Napel S, Nioche C, Orlhac F, Pati S, Pfaehler EAG, Rahmim A, Rao AUK, Scherer J, Siddique MM, Sijtsema NM, Socarras Fernandez J, Spezi E, Steenbakkers RJHM, Tanadini-Lang S, Thorwarth D, Troost EGC, Upadhaya T, Valentini V, van Dijk LV, van Griethuysen J, van Velden FHP, Whybra P, Richter C and Löck S. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology 2020; 295: 328-338.

[58] Sehn LH and Gascoyne RD. Diffuse large B-cell lymphoma: optimizing outcome in the context of clinical and biologic heterogeneity. Blood 2015; 125: 22-32.

[59] Thieblemont C, Tilly H, Gomes da Silva M, Casasnovas RO, Fruchart C, Morschhauser F, Haioun C, Lazarovici J, Grosicka A, Perrot A, Trotman J, Sebban C, Caballero D, Greil R, van Eygen K, Cohen AM, Gonzalez H, Bouabdallah R, Oberic L, Corront B, Choufi B, Lopez-Guillermo A, Catalano J, Van Hoof A, Briere J, Cabeçadas J, Salles G, Gaulard P, Bosly A and Coiffier B. Lenalidomide maintenance compared with placebo in responding elderly patients with diffuse large B-cell lymphoma treated with first-line rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. J Clin Oncol 2017; 35: 2473-2481.

[60] Johnson NA, Savage KJ, Ludkovski O, Ben-Neriah S, Woods R, Steidl C, Dyer MJ, Siebert R, Kuruvilla J, Klasa R, Connors JM, Gascoyne RD and Horsman DE. Lymphomas with concurrent BCL2 and MYC translocations: the critical factors associated with survival. Blood 2009; 114: 2273-2279.

[61] Pedersen M, Gang AO, Poulsen TS, Knudsen H, Lauritzen AF, Nielsen SL, Klausen TW and Nørgaard P. MYC translocation partner gene determines survival of patients with large B-cell lymphoma with MYC- or double-hit MYC/BCL2 translocations. Eur J Haematol 2014; 92: 42-48.

[62] Savage KJ, Johnson NA, Ben-Neriah S, Connors JM, Sehn LH, Farinha P, Horsman DE and Gascoyne RD. MYC gene rearrangements are associated with a poor prognosis in diffuse large B-cell lymphoma patients treated with R-CHOP chemotherapy. Blood 2009; 114: 3533-3537.

[63] Huang S, Nong L, Wang W, Liang L, Zheng Y, Liu J, Li D, Li X, Zhang B and Li T. Prognostic impact of diffuse large B-cell lymphoma with extra copies of MYC, BCL2 and/or BCL6: comparison with double/triple hit lymphoma and double expressor lymphoma. Diagn Pathol 2019; 14: 81.

[64] Willenbacher E, Willenbacher W, Weger R, Dominik W, Manzl C and Brunner A. Patients with double/triple copy number gains on C-MYC, BCL2, and/or BCL6 treated with standard chemotherapy have a similarly poor prognosis than those with high-grade B cell lymphoma with C-MYC and BCL2 and/or BCL6 rearrangements: a single-center experience on a consecutive cohort of large B cell lymphomas. Ann Hematol 2020; 99: 2125-2132.

[65] Jha AK, Mithun S, Purandare NC, Kumar R, Rangarajan V, Wee L and Dekker A. Radiomics: a quantitative imaging biomarker in precision oncology. Nucl Med Commun 2022; 43: 483-493.

[66] Ferrer Lores B, Mayorga-Ruiz I, Alberich Bayarri A, Morello-González D, Pastor-Galán I, Navarro-Cubells B, Serrano A, Teruel AI, Dosdá-Muñoz R, Solano C, Marti-Bonmati L and Terol MJ. Prognostic value of radiomics signature by diagnostic ^{18}F-FDG PET/CT analysis in aggressive non-Hodgkin's lymphoma. Blood 2018; 132: 1703-1703.

[67] Coskun N, Okudan B, Uncu D and Kitapci MT. Baseline 18F-FDG PET textural features as predictors of response to chemotherapy in diffuse large B-cell lymphoma. Nucl Med Commun 2021; 42: 1227-1232.

[68] Eertink JJ, Zwezerijnen GJC, Wiegers SE, Chamuleau MED, Lugtenburg P, De Jong D, Ylstra B, Mendeville M, Dührsen U, Hüttmann A, Schmitz C, Richter J, Klapper W, Hoekstra OS, de Vet HCW, Boellaard R and Zijlstra JM. Combining PET radiomic features with MYC gene rearrangement results in high prediction of outcome in diffuse large B-cell lymphoma. Blood 2021; 138: 451.

[69] Pyradiomics. [Accessed November 2022]; Available from: https://pyradiomics.readthedocs.io/en/latest/features.html#.

Table S1. Clinical and genomic variables according to response to treatment

Variables	Responders (*n* = 23)	Non-responders (*n* = 10)	Total (*N* = 33)	*P*-value
Clinical features, *n* (%)				
Sex				
Male	12 (52.17)	5 (50)	17 (51.52)	1.0
Female	11 (47.82)	5 (50)	16 (48.48)	
B symptoms				
Yes	11 (47.82)	5 (50)	16 (48.48)	1.0
No	12 (52.17)	5 (50)	17 (51.52)	
Bulky disease				
Yes	12 (52.17)	7 (70)	19 (57.58)	0.497
No	11 (47.82)	3 (30)	14 (42.42)	
R-IPI				
Very good	7 (30.43)	7 (70)	14 (42.42)	0.086
Good	13 (56.52)	3 (30)	16 (48.48)	
Poor	3 (13.04)	0 (0)	3 (9.09)	
Genomic features, *n* (%)				
MYC/8q24 rearrangement				
Yes	2 (8.70)	0 (0)	2 (6.06)	-
No	21 (91.30)	10 (100)	31 (93.94)	
MYC/8q24 amplification				
Yes	5 (21.74)	4 (40)	9 (27.27)	0.400
No	18 (78.26)	6 (60)	24 (72.73)	
BCL6/3q27 rearrangement				
Yes	5 (21.74)	3 (30)	8 (24.24)	0.673
No	18 (78.26)	7 (70)	25 (75.76)	
BCL6/3q27 amplification				
Yes	5 (21.74)	6 (60)	11 (33.33)	0.049
No	18 (78.26)	4 (40)	22 (67.67)	
BCL2/18q21 rearrangement				
Yes	7 (30.43)	0 (0)	7 (21.21)	0.073
No	16 (69.57)	10 (100)	26 (78.79)	
BCL2/18q21 amplification				
Yes	5 (21.74)	8 (80)	13 (39.39)	0.005
No	18 (78.26)	2 (20)	20 (60.61)	

R-IPI, Revised International Prognostic Index.

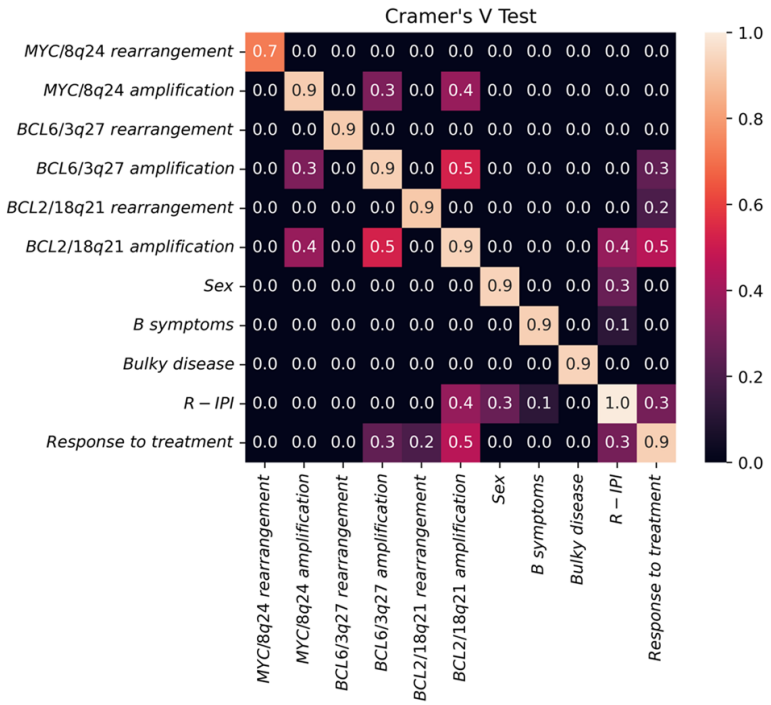

Figure S1. Association between clinical and genomic variables through Cramer's V test. R-IPI, Revised International Prognostic Index.

Table S2. Predictive models trained with clinical and radiomic variables and built with manual variable selection or by applying the linear discriminant analysis reduction method

Metrics	Predictive models (Logistic regression)	
	Manual selection	LDA
AIC	35.697	27.309
AUC	0.598	0.891
Sensitivity	69.6%	91.3%
Specificity	50.0%	70.0%
Balanced accuracy	59.8%	80.7%
Accuracy	63.6%	84.9%

AIC, Akaike Information Criterion; AUC, Area Under the Curve; LDA, Linear Discriminant Analysis.

Table S3. Predictive models trained with clinical, radiomic and genomic variables and built with manual variable selection or by applying the linear discriminant analysis reduction method

Metrics	Predictive models (Logistic regression)	
	Manual selection	LDA
AIC	34.816	27.451
AUC	0.757	0.904
Sensitivity	95.2%	100%
Specificity	50%	80%
Balanced accuracy	72.6%	90%
Accuracy	80.6%	93.9%

AIC, Akaike Information Criterion; AUC, Area Under the Curve; LDA, Linear Discriminant Analysis.