International Journal of Laboratory Hematology

WILEY

ISLH International Journal of Laboratory Hematology

ORIGINAL ARTICLE

Prediction of Lymphoma Aggressiveness Using Machine Learning Algorithms

Julien Cabo[1] | Benoît Bihin[2] | Nicolas Debortoli[3] | Virgine Lepage[1] | Reza Soleimani[1] | Rhita Bennis[1] | Julien Favressed[4] | Thierry Vander Borght[5] | Carlos Graux[6] | Caroline Fervaille[7] | Jonathan Degosserie[3] | Marie Pouplard[1] | François Mullier[1]

[1]Université Catholique de Louvain, CHU UCL Namur, Namur Thrombosis and Hemostasis Center (NTHC), Hematology Laboratory, Yvoir, Belgium | [2]Université Catholique de Louvain, CHU UCL Namur, Scientific Support Unit, Yvoir, Belgium | [3]Namur Molecular Tech (NMT), Yvoir, Belgium | [4]Université de Namur, Department of Pharmacy, Namur Thrombosis and Hemostasis Center (NTHC), Namur Research Institute for Life Sciences (NARILIS), Namur, Belgium | [5]Université Catholique de Louvain, CHU UCL Namur, Namur Thrombosis and Hemostasis Center (NTHC), Nuclear Medicine, Yvoir, Belgium | [6]Université Catholique de Louvain, CHU UCL Namur, Namur Thrombosis and Hemostasis Center (NTHC), hematology, Yvoir, Belgium | [7]Université Catholique de Louvain, CHU UCL Namur, Namur Thrombosis and Hemostasis Center (NTHC), Anatomopathology, Yvoir, Belgium

Correspondence: Julien Cabo (juliencabo1109@gmail.com)

Received: 6 September 2024 | **Revised:** 27 January 2025 | **Accepted:** 21 March 2025

Funding: The authors received no specific funding for this work.

Keywords: aggressive lymphoma | ensemble learning algorithms | lymph node cytology | machine learning | predictive modeling

ABSTRACT

Introduction: Lymph nodes are essential to diagnose lymphoid neoplasms, metastases, and infections. Some lymphomas, particularly aggressive non-Hodgkin lymphomas (NHL), need urgent diagnosis. Combining lymph node cytology (LNC) and flow cytometry (FC) with other rapidly available parameters through multivariable predictive models could offer valuable diagnostic information while waiting for anatomopathological results.

Materials and Methods: Results of 196 lymph node specimens were retrospectively analyzed for parameters like age, sex, LNC, FC, positron emission tomography scan, lymphocytosis, leukocytosis, lactate dehydrogenase (LDH) levels, and hemoglobin. We constructed five multivariable models predicting the aggressive nature of lymphoma as defined by the anatomopathological diagnostic. The first three were logistic regression models based on two (model 1), four (model 2), and up to 16 independent variables (model 3). The last two models were based on ensemble learning algorithms, bagging (model 4) and boosting (model 5), respectively. The performance of these five models was compared after 10-fold cross-validation, evaluating metrics such as sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC).

Results: Compared to individual variables associated with the aggressive nature of the lymphoma (AUCs from 0.69 to 0.87), the multivariable models achieved better AUCs, ranging from 0.88 to 0.94. The best model (model 5) achieved a sensitivity and a specificity of 77% and 94%, respectively.

Conclusion: LNC, FC, and other rapidly available parameters are associated with the aggressive nature of the lymphomas. It is possible to combine them in multivariable models to obtain a valuable diagnostic information and to initiate a prompt treatment.

Benoît Bihin, Nicolas Debortoli, Jonathan Degosserie, Marie Pouplard contributed equally to this work.

1 | Introduction

Lymph nodes are peripheral lymphoid tissues that play a role in the adaptive immune system. They can also be the origin of lymphoid neoplasms, the location for metastases from non-hematological tumors, as well as the site of infection [1]. The examination of lymph nodes is therefore of importance in clinical practice. Among the various lymphomas that may originate from lymph nodes, some require urgent diagnosis and are classically known as aggressive non-Hodgkin's lymphomas (NHL), either from B or T-cell lineages (see Table S1 in the Supporting Information for a representative, but not a comprehensive list) [2–6].

NHL is the most common hematological malignancy worldwide and accounts for 2.8% of cancer diagnoses and deaths [7]. Aggressive lymphomas account for a substantial proportion. In 2016, diffuse large B-cell lymphomas (DLBCL), Burkitt lymphoma (BL), and peripheral T-cell lymphoma (PTCL) were the most common aggressive NHLs, representing approximately 24% of all lymphoid malignancies diagnosed in the United States [8, 9]. Even more alarming, 25% were aggressive lymphomas in the age group of 15–39 years, climbing to 37% when including precursor B- and T-cell lymphoblastic lymphomas [9]. In addition, DLBCL was also identified as the most frequent aggressive B-cell lymphoma, accounting for approximately 30% of NHL adult cases in Western countries [1]. In aggressive B-cell lymphomas, about half of the patients present with advanced disease at the time of diagnosis and various clinical presentations [3]. Urgent treatment is usually required in cases of high proliferation rate of the tumor with bulky disease, end-organ damage, respiratory distress, obstructive symptoms, bone marrow involvement, and/or lysis syndrome [10]. Once the diagnosis of an aggressive lymphoma is confirmed, it is crucial to start the treatment without any delay.

Clinicians often consider cyto- and histopathological analyses as the gold standard for the investigation of lymph node neoplasms [11]. However, these analyses require sufficient materials to have thin slices and have a long turnaround time. Before obtaining anatomopathological results, clinical data along with various laboratory data are available, including lymph node cytology (LNC), flow cytometry (FC), and lactate dehydrogenase (LDH) levels. These laboratory analyses have a turnaround time inferior to 4 h. Positron emission tomography scan (PET-scan) data, especially the maximum standardized uptake value (SUVmax) which compensates for the variation in the amount of injected 2-deoxy-2-[18F]fluoro-D-glucose, is often also available before the lymph node sampling [12]. These preliminary results can provide valuable information for the diagnosis of lymphomas while waiting for the anatomopathological conclusion.

According to the Groupe Francophone d'Hématologie Cellulaire (GFHC), LNC is the microscopic examination of fine-needle aspiration or biopsy imprints, stained using the May-Grünwald Giemsa method [13, 14]. LNC is effective in identifying reactive processes, invasion of extra-hematological cells and is crucial for the characterization of lymphoid neoplasms due to the distinct morphological characteristics of lymphomatous cells [1, 15]. FC provides valuable information by evaluating the size of cells (forward scatter), their complexity (side scatter), the type of lymphoid B- and T-cell antigens, and the clonality of light chains [16]. The combination of LNC and FC enables the diagnosis of lymphoma with high sensitivity and specificity [16]. Certain morphological features help distinguish aggressive from non-aggressive lymphomas. Aggressive lymphomas have larger, irregular cells with prominent nucleoli, anisocytosis, and anisokaryosis, suggesting a fast-growing tumor requiring aggressive treatment [1, 15]. Non-aggressive lymphomas have smaller, uniform cells without nucleoli [15, 16]. In their review study, Cozzolino et al. reported that combining LNC and FC can effectively distinguish between reactive and lymphomatous lymph node specimens, with sensitivity and specificity ranging from 75% to 100% [16]. However, these researchers did not extend to differentiating between aggressive and indolent forms of lymphoma, nor did they include the potential of other parameters. There is also a lack of evidence in how LNC and FC can be beneficial for clinicians and anatomopathologists [17]. The complexity arises from the diverse patterns of morphological and immunophenotypic characteristics. Another concern is the between-observer subjectivity of LNC interpretation leading to inconsistencies in diagnoses. LDH levels and SUVmax are also relevant parameters to integrate in the diagnosis work-up. Elevated LDH levels can suggest an aggressive behavior of the lymphoma, correlating with the rate of cellular turnover [18]. It is the same for high SUV, with higher values often indicating more aggressive or transformed lymphoma [19, 20].

Building a model that predicts the aggressive or non-aggressive nature of lymphomas using LNC, FC, and other rapidly available parameters would therefore seem to be an appropriate way of substantially speeding up the time to diagnosis, especially as the rapid development of machine learning tools makes it possible to use ever greater quantities of data to improve predictive models [21].

Machine learning methods model the relationship between the lymphoma's aggressiveness and predictors. Logistic regression, for example, uses an equation with interpretable terms and can involve predefined predictors or a variable selection algorithm like stepwise forward selection.

The model fits the data by finding predictor coefficients that maximize the model's likelihood.

On the other hand, some machine learning methods predict a result not from a single model but from a set of models, most often decision trees. Ensemble learning algorithms include bagging [22] and boosting [23]. The random forest algorithm [24] is based on bagging, that is, the generation of a large number of bootstrap samples from the initial dataset, with each bootstrap sample including only some of the potential predictors. A decision tree is fitted to each of these samples. The random forest predictions are based on the most frequent prediction of this set of models. The second, eXtreme Gradient Boosting (XGBoost) model [25] is based on boosting, each of which corrects the errors made by the previous model. The individual predictions of each model are then combined to form the final prediction.

The aim of this study was to determine to what extent it is possible to predict the aggressive or non-aggressive nature of

lymphoma as defined by anatomopathological examination by combining the information provided by LNC, FC, and other readily available parameters. To do this, we conducted a retrospective single-center study in which we compared the predictive performance of five types of models, ranging from the simplest—a pre-specified logistic regression model based on LNC and FC—to the most complex—a set of models constructed by gradient boosting integrating the results of 16 laboratory variables.

2 | Materials and Methods

2.1 | Design of the Study

The investigation was carried out in the hematology laboratory at CHU UCL Namur (Yvoir, Belgium) and was based on the retrospective analysis of lymph node specimens collected between 2020 and 2022. In total, 196 specimens were evaluated for various parameters, including sex, age, history, SUVmax, lymphocytosis, leukocytosis, LDH, C-reactive protein, hemoglobin, aggressive morphological characteristics by LNC (see examples in Supporting Information), forward scatter ratio (FSR) by FC, and conclusions from anatomopathological protocols. Out of the 196 specimens, 34 specimens were considered unsuitable for inclusion in the study due to insufficient material for comprehensive analysis. The definitive categorization of each specimen was determined from the anatomopathological reports, splitting them into two principal classifications for evaluation: aggressive lymphomas versus the other category comprising indolent lymphomas, reactive lymphadenopathy, and extra-hematological malignancies. In the aggressive category, lymphomas were classified respectively into de novo DLBCL ($n=31$), relapsed DLBCL ($n=4$), transformation from indolent B-cell lymphomas ($n=3$), high-grade B-cell lymphomas ($n=4$), BL ($n=1$), aggressive B lymphoma not otherwise specified ($n=1$), mantle cell lymphoma (MCL), blastoid variant ($n=1$) and T-lymphoblastic lymphoma ($n=1$). Furthermore, we included in this category one case of lymph node infiltration by multiple myeloma and one by myeloid sarcoma with immature cells. In a second analysis, MCL other than blastoid variant and grade 3a follicular lymphomas (FL) ($n=15$) were reclassified as aggressive lymphomas, as the clinical course and the treatment regimen are still controversial in the literature [26, 27].

2.2 | Parameters

The details about lymph node sampling, SUVmax, LDH level measurement, lymphocytosis, leukocytosis, hemoglobin measurements, LNC, and sample preparation for FC are available in the Supporting Information.

The FSR was calculated by dividing the average forward scatter value of monoclonal cells by the average forward scatter value of polyclonal cells, which included T and B lymphocytes constituting the normal cells of a lymph node environment. The calculation of this FSR implied that monoclonality was confirmed with fluorochrome-labeled antibodies, as evidenced by either light chain monoclonality in B-cell lymphomas or T-cell receptor monoclonality in T-cell lymphomas.

2.3 | Statistical Analysis

2.3.1 | Description of Sample Characteristics

Continuous variables were summarized by their median and interquartile range, while categorical variables were summarized by the number and percentage of each category. The area under the operating characteristic curve (AUC) with the 95% confidence interval (95% CI) was provided to measure the predictive performance of each potential predictor. The descriptive analysis (Table 1) presents the raw data. Values significantly higher than 0.5 suggest an association, with the strength of this association increasing as the AUC tends to 1, which is a perfect test [28]. For predictive model construction, missing values were replaced by the mode for categorical variables and the median for continuous variables. Continuous variables, except for age and hemoglobin concentration, were log-transformed.

2.3.2 | Characteristics of the Predictive Models

Five predictive models were developed to analyze the data and assess the effectiveness of various predictors.

Model 1: Logistic regression model based on two prespecified independent variables.

$$Y_i = \beta_0 + \beta_1 * LNC + \beta_2 * \mathrm{FSR} \quad (1)$$

Yi represents the logarithm of the odds that a lymph node is classified as aggressive. To clarify, if the probability of a lymph node being aggressive is 50%, the odds are 1, and the log odds (Yi) are 0. β_0 represents the log odds of a lymph node being aggressive when both the LNC and FSR are at baseline levels (i.e., LNC = 0 and FSR ≤ 1). β_1 indicates how the log odds of aggressiveness change when the LNC increases from 0 to 1. Specifically, exp. (β_1) represents the odds ratio, which quantifies the change in the odds of aggressiveness for a lymph node with LNC = 1 compared with one with LNC = 0. β_2 reflects how the log odds of aggressiveness change when the FSR increases from a value of 1 to a value greater than 1. It captures the effect of the FSR on the likelihood of aggressiveness. The equation can be utilized to measure the association between each predictor (LNC and FSR) and the odds of lymph node aggressiveness as well as to predict the probability of aggressiveness based on the predictors. This is done using the following transformation:

$$\mathrm{Probability} = \frac{e^{Y_i}}{1 + e^{Y_i}} \quad (2)$$

This equation converts the log odds into a probability that ranges from 0 to 1, providing a clear interpretation of the likelihood that a lymph node is aggressive based on its characteristics.

Model 2: Logistic regression model based on four prespecified independent variables.

$$Y_i = \beta_0 + \beta_1 * \mathrm{LNC} + \beta_2 * \mathrm{FSR} + \beta_3 * \mathrm{SUVmax} + \beta_4 * \mathrm{LDH} \quad (3)$$

TABLE 1 | Baseline characteristics according to the anatomopathological final category.

	Anatomopathological category			
Characteristic	**Overall (*N* = 162)**	**Other categories (*N* = 113)**	**Aggressive lymphomas (*N* = 49)**	**AUC [95% CI]**
Sex: female	67 (41%)	48 (42%)	19 (39%)	0.52 [0.44–0.60]
Age	63 (46, 72)	58 (42, 69)	70 (59, 77)	**0.69 [0.61–0.78]**
Adenopathy	126/139 (91%)	88/106 (93%)	38/43 (86%)	0.53 [0.47–0.59]
Splenomegaly	20/85 (24%)	14/58 (24%)	6/27 (22%)	0.51 [0.41–0.61]
Hepatomegaly	2/76 (2.6%)	2/51 (3.9%)	0/25 (0%)	0.52 [0.49–0.55]
B symptoms	17/79 (22%)	12/53 (23%)	5/26 (19%)	0.52 [0.42–0.61]
SUVmax	11 (7, 14)	11 (6, 11)	15 (11, 23)	**0.74 [0.65–0.82]**
Leucocytosis	7.3 (5.6, 9.9)	7.3 (6.0, 10.2)	6.3 (4.8, 9.7)	**0.60 [0.50–0.70]**
Lymphocytosis	1.4 (1.0, 1.9)	1.4 (1.2, 2.3)	1.3 (0.8, 1.5)	**0.64 [0.55–0.73]**
LDH	216 (182, 275)	204 (177, 236)	254 (222, 406)	**0.74 [0.66–0.83]**
CRP	9 (2, 33)	8 (1, 26)	18 (3, 46)	0.59 [0.49–0.70]
Hemoglobin	13.0 (10.9, 14.1)	13.0 (10.9, 14.4)	12.8 (10.9, 14.1)	0.53 [0.43–0.64]
Positive AEHM	8 (4.9%)	6 (5.3%)	2 (4.1%)	0.51 [0.47–0.54]
Positive AHM	65 (40%)	36 (32%)	29 (59%)	**0.64 [0.55–0.72]**
LNC aggressive morphology	45 (28%)	10 (8.8%)	35 (71%)	**0.81 [0.74–0.88]**
FSR	1.00 (1.00, 1.24)	1.00 (1.00, 1.01)	1.35 (1.19, 1.49)	**0.87 [0.80–0.94]**

Note: The results are indicated either as a percentage, which reflects the frequency of occurrence relative to the number of samples with available data (%), or as a median value accompanied by the interquartile range, indicating the 25th and 75th percentile values. Bold values are statistically significant (confidence intervals not crossing 0.5).
Abbreviations: AEHM, antecedents of an extra-hematological malignancy; AHM, antecedent of hematological malignancy; AUC, area under the receiver operating characteristic curve; CI, confidence interval; CRP, C-reactive protein; FSR, forward scatter ratio; LDH, lactate dehydrogenase; LNC, lymph node cytology; SUVmax, standard uptake value maximum.

Model 3: logistic regression model based on automatic variable selection based on the Akaike information criterion [29]. The 16 potential predictors from Table 1 were included.

For the first three models, the FSR variable was dichotomized (FSR = 1 if FSR > 1, otherwise FSR = 0) because logistic regression assumes a linear relationship between the independent variable and the response, which was difficult to support given the significant asymmetry in the distribution of FSR values. LNC is 1 if morphological orients to aggressive lymphoma and 0 otherwise.

Model 4: Random forest based on the 16 predictors, using the randomForest function from the randomForest package with default parameters. The relative importance of the predictors is measured using the Gini index given by the importance function of the randomForest package.

Model 5: A set of models built from the XGBoost algorithm, using the XGBoost function from the XGBoost package with the following parameters: the maximum depth of a tree is fixed to 2, the objective is logistic regression for binary classification. The relative importance of the predictors in these models is measured using the gain calculated using the xgb.importance function in the XGBoost package.

2.3.3 | Comparison of the Predictive Performance of the Models

Evaluation of the predictive performance of these models was based on 6 indices: accuracy, AUC, positive predictive value (PPV), negative predictive value (NPV), sensibility, and specificity. For comparison, the indices corresponding to the FSR and LNC variables taken in isolation were also presented. To mitigate the risk of model overfitting, particularly for the potentially more complex models 3, 4, and 5, all model performance metrics were subjected to 10-fold cross-validation. The model, based on the full dataset of 162 lymph nodes (Figure 2), was evaluated by dividing the data into 10 parts. In each iteration, 9 parts were used for training and 1 for testing, with performance metrics averaged over 10 iterations. The final statistics, detailed in Table 3, represent these averaged results. All analyses were performed using R 4.2.2 (The R Foundation for Statistical Computing, Austria, Vienna, 2022) and the following packages: caret [30], randomForest [24], XGBoost [25].

3 | Results

3.1 | Lymph Node Characteristics

The analysis of the clinical and laboratory characteristics among 162 lymph node specimens, categorized into two groups: "Aggressive lymphomas" ($n = 49$) and "Other categories" ($n = 113$), is presented in Table 1. Gender was slightly imbalanced, with females accounting for 41% of the overall population, but was not a valuable predictor (AUC = 0.52, CI 95: 0.44–0.60). The difference was more pronounced for age, with the median age significantly higher in the aggressive lymphoma group (70 years) compared with the other categories (58 years), with an AUC of 0.69 (CI 95: 0.61–0.78). Higher SUVmax and LDH levels were observed in aggressive cases, each marked by an AUC of 0.74 (CI 95: 0.65–0.82 and 0.66–0.83 respectively). Lymphocytosis and leukocytosis demonstrated moderate predictive power, with AUCs of 0.64 (CI 95: 0.55–0.73) and 0.60 (CI 95: 0.50–0.70). C-reactive protein and hemoglobin levels had insignificant predictive power, with their 95% CI crossing 0.50. A history of hematological cancer was somewhat predictive (AUC: 0.64, CI 95: 0.55–0.72), unlike a history of other cancers (AUC: 0.51, CI 95: 0.47–0.54). Most notably, aggressive morphological characteristics obtained with LNC analysis and elevated FSR showed significant differences, being considerably higher in aggressive lymphomas (71% for morphology, and a median FSR of 1.35 in aggressive lymphomas versus 1.00 in other categories). These findings were corroborated by their high AUC values (0.81, CI 95: 0.74–0.88 for morphology and 0.87, CI 95: 0.80–0.94 for FSR), indicating their strong predictive value. This analysis underlines the importance of a multiparametric diagnostic approach in lymphoma classification, with certain characteristics like age, SUVmax, LDH, antecedents of hematological malignancy, LNC, and FSR demonstrating predictive capabilities.

3.2 | Description of Model Characteristics

In this section, the characteristics of the five predictive models have been compared. The first three are logistic regression models, while the next two use ensemble learning algorithms.

3.2.1 | Logistic Regression–Based Models

Model 1 was a pre-specified model including only FSR and LNC. In this model, an FSR > 1 and a positive LNC increased the odds of the lymph node being classified as aggressive at anatomopathological examination by factors of 11.5 (CI 95: 4.3–35.4) and 17.5 (CI 95: 6.6–51.4), respectively (Table 2). According to this model, a patient with a positive LNC and an FSR > 1 would have an 89% chance of having an aggressive lymphoma (Figure 1A).

Model 2 was a pre-specified model that includes FSR, LNC, SUVmax, and LDH. In this model, an FSR > 1, a positive LNC, an increase of one log of SUVmax, or LDH increased the odds of the lymph node being classified as aggressive at anatomopathological examination by factors of 8.6 (CI 95: 3.0–27.4), 13.4 (CI 95: 4.7–42.3), 3.1 (CI 95: 1.1–9.9), and 3.5 (CI 95: 1.2–12.2), respectively (Table 2). According to this model, a patient with a positive LNC and a FSR > 1 would have a 37%–99% chance of having an aggressive lymphoma based on SUVmax and LDH levels (Figure 1B).

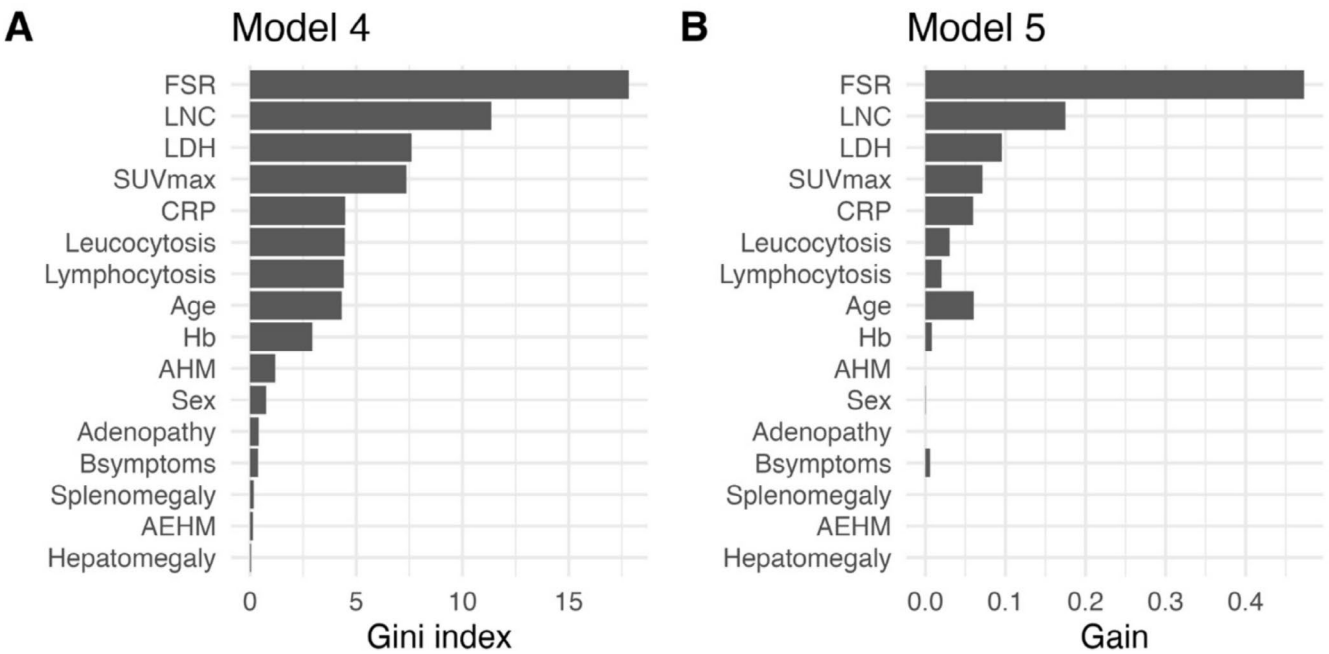

FIGURE 1 | Relative importance of the predictors. (A) Relative importance of the predictors included in the model 4, assessed by the Gini index. (B) Relative importance of the predictors included in the model 5, assessed by the gain associated with each variable. Abbreviations: FSR, forward scatter ratio; LNC, lymph node cytology; LDH, lactate dehydrogenase; SUVmax, standard uptake value maximum; CRP, C-reactive protein; Hb, hemoglobin; AHM, antecedent of hematological malignancy; AEHM, antecedents of an extra-hematological malignancy.

TABLE 2 | Logistic regression model coefficients. The odds ratio of a predictor and its 95% confidence interval are obtained by exponentiating the table values.

	Model 1		Model 2		Model 3	
Characteristic	**Est.**	**95% CI**	**Est.**	**95% CI**	**Est.**	**95% CI**
(Intercept)	−3.26	−4.35, −2.39	−12.6	−20.0, −6.42	−9.31	−14.0, −5.42
FSR > 1	2.44	1.45, 3.57	2.15	1.10, 3.31	2.84	1.63, 4.25
LNC aggressive	2.86	1.89, 3.94	2.60	1.54, 3.74	3.33	2.09, 4.83
SUVmax			1.14	0.114, 2.29	1.50	0.390, 2.73
LDH			1.24	0.155, 2.50		
CRP					0.695	0.306, 1.16
B symptoms					−1.94	−3.96, −0.084
Age					0.034	−0.001, 0.073
Adenopathy					−1.56	−3.51, 0.358

Abbreviations: 95% CI, 95% confidence interval; CRP, C-reactive protein; Est., estimate of the beta coefficients of the model; FSR, forward scatter ratio; LDH, lactate dehydrogenase; LNC, lymph node cytology; NA, not applicable; SUVmax, standard uptake value maximum.

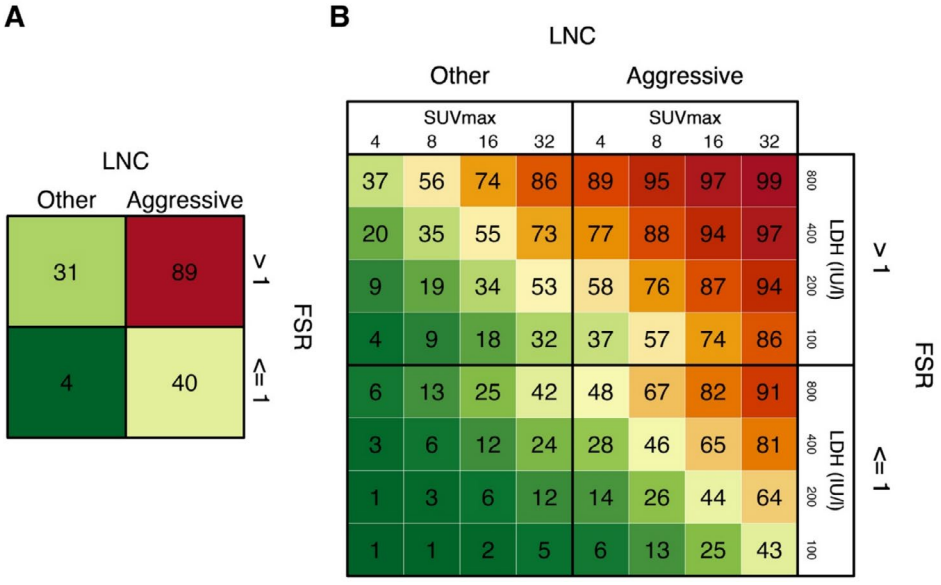

FIGURE 2 | Probability of a lymph node being considered aggressive compared with anatomopathological examination according to models 1 (A) and 2 (B). For example, according to model 1, a lymph node with no aggressive patterns in LNC examination and an FSR > 1 would have a 31% chance of being considered aggressive on anatomopathological examination. According to model 2, a lymph node with no aggressive patterns in LNC examination, an FSR > 1, a SUVmax of 8, and LDH levels of 400 IU/L would have a 35% chance of being considered aggressive on anatomopathological examination. Abbreviations: SUVmax, standard uptake value maximum; LDH, lactate dehydrogenase; LNC, lymph node cytology; FSR, forward scatter ratio.

Model 3 automatically selected predictors from a list of 16 potential predictors presented in Table 1. This model included the following eight variables: LNC, FSR, CRP, B symptoms, SUVmax, age, and adenopathy. The model coefficients are presented in Table 2. With eight variables, a graphical representation of the model prediction would be difficult to read.

3.2.2 | Ensemble Learning–Based Models

Model 4 was a random forest based on the 16 variables presented in Table 1. This forest was made up of 500 decision trees, each built on a sample constructed by bootstrapping from the original sample and randomly drawing 4 independent variables from the 16 potential ones. One way of comparing the relative weights of the independent variables was to measure the Gini index [31]. The four most important variables in this model were FSR, LNC, LDH, and SUVmax. These were followed by CRP, leukocytosis, lymphocytosis, age, and, to a lesser extent, hemoglobin (Figure 2A).

Model 5 was an ensemble of 25 decision trees with no more than two splits constructed using the XGBoost algorithm, within which each new model attempts to correct the errors made by the previous model. The relative importance of each of the 16

TABLE 3 | Comparison of model prediction performance. Shaded lines correspond to performance indexes measured on isolated variables (FSR and LNC).

	Variables	Accuracy	AUC	NPV	PPV	Sensitivity	Specificity
FSR > 1	1	0.78	0.80*	0.92	0.59	0.86	0.74
LNC aggressive morphology	1	0.85	0.81	0.88	0.78	0.71	0.91
Model 1	2	0.85	0.92	0.85	0.89	0.61	0.95
Model 2	4	0.88	0.92	0.88	0.89	0.72	0.94
Model 3	16	0.85	0.88	0.88	0.76	0.74	0.89
Model 4	16	0.85	0.94	0.89	0.74	0.69	0.91
Model 5	16	0.89	0.93	0.90	0.91	0.77	0.94

Abbreviations: AUC, area under the receiver operating characteristic curve; FSR, forward scatter ratio; LNC, lymph node cytology; NPV, negative predictive value; PPV, positive predictive value.
*The AUC for the FSR not dichotomized is 0.87.

independent variables potentially included in each tree was measured by the gain associated with each variable. The most important variables in model 5 were FSR, LNC, LDH, and SUVmax, followed by age, CRP, and, to a lesser extent, leukocytosis and lymphocytosis (Figure 2B).

3.2.3 | Comparison of Model Predictive Performance

In terms of AUC, all the multi-variable models appeared to be more discriminating than LNC or FSR alone, with the exception of model 3 (logistic regression with automatic variable selection) whose performance (AUC = 0.88) was only slightly better than FSR alone when it was not dichotomized (Table 3). Model 4, based on the random forest algorithm, showed the highest AUC (0.94), followed closely by model 5 (XGBoost model) with an AUC of 0.93.

In terms of accuracy, only two models exceeded the performance of diagnosis based on LNC alone, namely model 2 (logistic regression model based on FSR, LNC, SUVmax, and LDH) and model 5 (XGBoost model) with, respectively, an accuracy of 88% and 89%. This translated into slightly higher sensitivities and specificities than LNC alone (71% and 91%), namely 72% and 94% for model 2 and 77% and 94% for model 5, respectively (Table 3).

4 | Discussion

The aim of this study was to determine to what extent models based on LNC, FSR, and incorporating other rapidly available parameters can predict the outcome of the anatomopathological examination, in this case, the aggressive nature of the lymphomas.

We found that, of the 16 predictive variables studied, 8 appeared to be more or less associated with the aggressive nature of the lymphoma, namely FSR (AUC = 0.87, CI 95: 0.80–0.94), LNC (AUC = 0.81, CI 95: 0.74–0.88), SUVmax (AUC = 0.74, CI 95: 0.65–0.82), LDH (AUC = 0.74, CI 95: 0.66–0.83), age (AUC = 0.69, CI 95: 0.61–0.78), AHM (AUC = 0.64, CI 95: 0.55–0.72), lymphocytosis (AUC = 0.64, CI 95: 0.55–0.73) and, to a lesser extent, leukocytosis (AUC = 0.60, CI 95: 0.50–0.70) (Table 1). This corroborates our initial hypothesis that FSR, LNC, SUVmax, and LDH play a dominant role in these predictive models. This observation also emerges from the indirect measures of the relative importance of the variables in the random forest and in the XGBoost model (Figure 2).

The FSR alone provided important information, with an accuracy of 0.78 and an AUC of 0.87. The LNC alone achieved accuracy, sensitivity, and specificity of 0.85, 0.71, and 0.91, respectively. With regard to the multivariate models, three of the five models did not achieve better accuracy than LNC alone: the logistic regression model based on LNC and FSR (model 1), the logistic regression model with automatic variable selection (model 3), and the model based on the random forest (model 4). The performance of model 1 (accuracy = 0.85, AUC = 0.92) suggested that simply combining LNC and FSR in a logistic regression model was not sufficient to produce a model that performs satisfactorily on a sample that has not been used to train it. While there was an improvement in the AUC, this did not translate into improved accuracy. The same applied to model 3 (accuracy = 0.85, AUC = 0.88). In this case, the logistic regression with automatic variable selection might have been overfitted to the specificities of the training set, resulting in lower performance on the test set. In the case of the random forest (model 4), the low accuracy (0.85) can perhaps be explained by the fact that the trees constructed from variables that had little to do with the aggressive nature of the lymphoma diluted the overall predictive performance. Paradoxically, this model had the highest AUC (0.94), indicating that it has a 94% chance of correctly identifying an aggressive lymphoma.

On the other hand, our data highlighted an interest in the prespecified logistic regression model based on LNC, FSR, SUVmax, and LDH (model 2) as well as the XGBoost model (model 5). The logistic regression model's effectiveness can be attributed to the selection of 4 predefined variables, which were strongly associated with the aggressive nature of the lymphoma. In the case of the XGBoost model, it was based on a relatively small number of successive trees (25) and on trees of low complexity (maximum two levels) enabling

interpretability despite boosting model black boxes. These two models may have benefited from a good ratio between the richness of the variables included and the relatively limited complexity of the model. It is possible that on larger datasets, more complex algorithms relying on black boxes (e.g., boosting and neural network models) could outperform these two models. Machine learning has recently emerged in laboratory medicine as a beneficial tool in vast datasets with a broad range of applications such as detection of pre-analytical errors, disease and outcome prediction, or image classification [21]. In the current clinical workflow, treatment for aggressive lymphomas is typically initiated only after anatomopathological confirmation, which can delay urgent therapy in cases of high-grade disease for 24–48 h. Our algorithm (model 5) could enhance earlier identification of aggressive cases based on readily available parameters. Specifically, the sensitivity of 0.77 indicates that 77% of aggressive cases could potentially be flagged and treated earlier, reducing the time to intervention. Meanwhile, the high specificity of 0.94 ensures that the algorithm minimizes the risk of overtreatment by correctly excluding the majority of non-aggressive cases. While the algorithm does not replace anatomopathological confirmation, its use as a preliminary diagnostic aid could significantly enhance clinical decision-making, particularly in high-risk cases where early treatment is critical.

Although our study showed promising results, there are certain limitations. Primarily, the dataset utilized is small in scale, which may limit the generalizability of our models. This suggests that the models may be subject to some overfitting, meaning they have adapted too closely to the specific characteristics of the training data at the cost of their ability to generalize to new datasets [21, 32]. To address the overfitting issue, the dataset should be increased to lower the sampling noise and to introduce cross-validation with supplementary/external datasets [21, 32]. Second, the diversity of malignancies included in our dataset reflects the variety of cases encountered in clinical practice. However, this heterogeneity introduces variability in clinical presentations, natural histories, and responses to treatment, which could impact the generalizability of the findings. In addition, future studies could refine the model to allow discrimination among all categories (reactive, non-aggressive, aggressive, and extra-hematological). It also depends on the knowledge of experts in laboratory medicine and anatomopathology to correctly identify aggressive lymphoma cells and confirm monoclonality using FC. This usually necessitates that at least one senior specialist in laboratory medicine, and preferably two distinct experts, read the imprint smears [33]. In the 2019 GFHC survey, 3 specialists in laboratory medicine out of 39 expressed reservations, viewing LNC as less valuable than histological reporting and arguing that young biologists lack training nowadays [33]. Another limitation of this study is that the anatomopathological diagnoses were not independently reviewed by the research team. However, all cases were double-checked by specialists whenever doubts arose. Furthermore, there are different sampling methods for LNC, such as fine-needle lymph node aspiration, biopsy, and cytospins, which can lead to different conclusions [33]. Additionally, the preparation method, such as imprint or smear techniques, can also result in variations, and LNC is only performed on a fraction of the lymph node specimen, not the entirety. Consequently, the findings depend on the lymph node tissue homogeneity, raising the possibility of a missed diagnosis if an unrepresentative section is analyzed, despite multiplying the imprints. In cases of sclerosis, the representativeness of the tumoral tissue on smears may be compromised, with potentially the target cells undetected [13]. The global architecture and representativity of the ganglion cannot be appreciated, contrary to anatomopathological examination. Surgeons frequently cut the specimens into multiple sections, sending the largest piece to the anatomopathology department while the rest is given to different laboratories for LNC, FC, and genetic analyses. This division introduces uncertainty regarding whether the same tissue is assessed in each analysis. FC has limitations in the identification of high-grade B-NHLs lacking surface light chain expression or anaplastic non-Hodgkin lymphomas ($N = 2$ in our study) [34, 35]. Previous studies have shown that necrosis and cell fragility can lead to poor detection by FC, potentially producing false negatives [36, 37]. Although light chain analysis and specific immunophenotyping generally provide high sensitivity in determining clonality, a small but distinct percentage of cases remain inconclusive across all reported studies [34, 35, 38, 39]. For instance, 13.8% (4/29) and 17.4% (4/23) were inconclusive for anaplastic large cell lymphoma [34, 39]. Similarly, 9.7% (29/299) were inconclusive [35] and 11.0% (50/456) of B-NHL cases analyzed by FC at Tokai University Hospital lacked surface immunoglobulin light chain expression [38].

It is also still unclear whether certain lymphomas should be considered aggressive or not. Regarding the FL, there is a distinction between FL3A and FL3B, with FL3B generally following a more aggressive course similar to DLBCL [6, 26]. The same issues are discussed about the MCL. Aggressive variants, on the other hand, progress rapidly with widespread lymph node involvement and extranodal tissues, blastoid or pleomorphic morphology, high Ki-67 proliferation, and high expression of tumor protein p53 [40]. In multiple logistic regression analysis, these two less-aggressive NHL closely approached the threshold for aggressive lymphomas. An integrated analysis that grouped FL3A and MCL with aggressive lymphomas gave a stable AUC of 0.94.

5 | Conclusion

FSR, LNC, and other parameters such as SUVmax and LDH are associated with the aggressive nature of lymphoma. A logistic regression model combining these four parameters predicts the outcome of the anatomopathological examination with a sensitivity of 72% and a specificity of 94%, which could have significant clinical implications. An XGBoost model based on 16 features reached 77% sensitivity and 94% specificity. Integrating more parameters on larger quantities of data using ensemble learning algorithms should make it possible to substantially improve these predictions.

Author Contributions

J.C., R.S., M.P., and F.M. concepted and designed the study. J.C., V.L., R.S., and R.B. acquired the data. J.C., B.B., N.D., J.F., and F.M. interpreted the data. J.C. drafted the manuscript. All authors (J.C., B.B., N.D., V.L., R.S., R.B., J.F., T.V.B., C.G., C.F., J.D., M.P., and F.M.) critically revised the manuscript, approved its final version, and agree to be accountable for all aspects of the presented work.

Ethics Statement

The authors have nothing to report.

Consent

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. C. E. Swerdlow, S. H. Harris, N. L. Jaffe, et al., eds., *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*, 4th ed. (IARC, 2017).

2. J. Zain and A. Kallam, "Challenges in Nodal Peripheral T-Cell Lymphomas: From Biological Advances to Clinical Applicability," *Frontiers in Oncology* 13, no. 2023 (2023): 1150715.

3. C. Beham-Schmid, "Aggressive Lymphoma 2016: Revision of the WHO Classification," *Memo-Magazine of European Medical Oncology* 10, no. 4 (2017): 248–254.

4. K. E. Grimm and D. P. O'Malley, "Aggressive B Cell Lymphomas in the 2017 Revised WHO Classification of Tumors of Hematopoietic and Lymphoid Tissues," *Annals of Diagnostic Pathology* 38 (2019): 6–10.

5. R. Alaggio, C. Amador, I. Anagnostopoulos, et al., "The 5th Edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms," *Leukemia* 36, no. 7 (2022): 1720–1748.

6. E. Campo, E. S. Jaffe, J. R. Cook, et al., "The International Consensus Classification of Mature Lymphoid Neoplasms: A Report From the Clinical Advisory Committee," *Blood, the Journal of the American Society of Hematology* 140, no. 11 (2022): 1229–1253.

7. F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA: A Cancer Journal for Clinicians* 68, no. 6 (2018): 394–424.

8. J. R. Cerhan, "Epidemiology of Aggressive Lymphomas," in *Aggressive Lymphomas* (Springer, 2019), 3–40.

9. K. A. Blum, F. G. Keller, S. Castellino, A. Phan, and C. R. Flowers, "Incidence and Outcomes of Lymphoid Malignancies in Adolescent and Young Adult Patients in the United States," *British Journal of Haematology* 183, no. 3 (2018): 385–399.

10. E. N. Mugnaini and N. Ghosh, "Lymphoma," *Primary Care; Clinics in Office Practice* 43, no. 4 (2016): 661–675.

11. S. H. Kroft, C. E. Sever, A. Bagg, et al., "Laboratory Workup of Lymphoma in Adults: Guideline From the American Society for Clinical Pathology and the College of American Pathologists," *American Journal of Clinical Pathology* 155, no. 1 (2021): 12–37.

12. C. Hoekstra, I. Paglianiti, O. Hoekstra, et al., "Monitoring Response to Therapy in Cancer Using [18 F]-2-Fluoro-2-Deoxy-D-Glucose and Positron Emission Tomography: An Overview of Different Analytical Methods," *European Journal of Nuclear Medicine* 27 (2000): 731–743.

13. J.-F. Lesevê, K. Maloum, C. Settegrana, et al., "Lecture D'une Cytologie Ganglionnaire: Proposition D'une Procédure Type," *Revue de Biologie Médicale* 357 (2020): 29–44.

14. E. Piaton, M. Fabre, I. Goubin-Versini, et al., "Guidelines for May-Grünwald–Giemsa Staining in Haematology and Non-Gynaecological Cytopathology: Recommendations of the French Society of Clinical Cytology (SFCC) and of the French Association for Quality Assurance in Anatomic and Cytologic Pathology (AFAQAP)," *Cytopathology* 27, no. 5 (2016): 359–368.

15. P. Felman and O. Gentilhomme, *Atlas de Cytopathologie Ganglionnaire* (Arnette, 1997).

16. I. Cozzolino, M. Rocco, G. Villani, and M. Picardi, "Lymph Node Fine-Needle Cytology of Non-Hodgkin Lymphoma: Diagnosis and Classification by Flow Cytometry," *Acta Cytologica* 60, no. 4 (2016): 302–314.

17. B. A. Meda, D. H. Buss, R. D. Woodruff, et al., "Diagnosis and Subclassification of Primary and Recurrent Lymphoma: The Usefulness and Limitations of Combined Fine-Needle Aspiration Cytomorphology and Flow Cytometry," *American Journal of Clinical Pathology* 113, no. 5 (2000): 688–699.

18. V. Jurisic, S. Radenkovic, and G. Konjevic, "The Actual Role of LDH as Tumor Marker, Biochemical and Clinical Aspects," *Advances in Cancer Biomarkers: From Biochemistry to Clinic for a Critical Revision* 867 (2015): 115–124.

19. G. Alobthani, V. Romanov, M. Kayako, K. Matsunaga, W. Tadashi, and T. Mitsuaki, "Value of F-FDG PET/CT in Discrimination Between Indolent and Aggressive Non-Hodgkin's Lymphoma: A Study of 328 Patients," *Hellenic Journal of Nuclear Medicine* 21, no. 1 (2018): 7–14.

20. A. Noy, H. Schöder, M. Gönen, et al., "The Majority of Transformed Lymphomas Have High Standardized Uptake Values (SUVs) on Positron Emission Tomography (PET) Scanning Similar to Diffuse Large B-Cell Lymphoma (DLBCL)," *Annals of Oncology* 20, no. 3 (2009): 508–512.

21. S. De Bruyne, M. M. Speeckaert, W. Van Biesen, and J. R. Delanghe, "Recent Evolutions of Machine Learning Applications in Clinical Laboratory Medicine," *Critical Reviews in Clinical Laboratory Sciences* 58, no. 2 (2021): 131–152.

22. L. Breiman, "Bagging Predictors," *Machine Learning* 24 (1996): 123–140.

23. J. Friedman, T. Hastie, and R. Tibshirani, "Additive Logistic Regression: A Statistical View of Boosting (With Discussion and a Rejoinder by the Authors)," *Annals of Statistics* 28, no. 2 (2000): 337–407.

24. A. Liaw and M. Wiener, "Classification and Regression by randomForest," *R News* 2, no. 3 (2002): 18–22.

25. T. Chen, T. He, M. Benesty, et al., "Xgboost: Extreme Gradient Boosting, R Package Version 0.4-2," (2015) 1–4.

26. K. Koch, E. Hoster, M. Ziepert, et al., "Clinical, Pathological and Genetic Features of Follicular Lymphoma Grade 3A: A Joint Analysis of the German Low-Grade and High-Grade Lymphoma Study Groups GLSG and DSHNHL," *Annals of Oncology* 27, no. 7 (2016): 1323–1329.

27. D. E. Issa, S. A. van de Schans, M. E. Chamuleau, et al., "Trends in Incidence, Treatment and Survival of Aggressive B-Cell Lymphoma in the Netherlands 1989–2010," *Haematologica* 100, no. 4 (2015): 525.

28. F. Habibzadeh, P. Habibzadeh, and M. Yadollahie, "On Determining the Most Appropriate Test Cut-Off Value: The Case of Tests With Continuous Results," *Biochemia Medica* 26, no. 3 (2016): 297–307.

29. H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control* 19, no. 6 (1974): 716–723.

30. M. Kuhn, "Building Predictive Models in R Using the Caret Package," *Journal of Statistical Software* 28 (2008): 1–26.

31. J. L. Gastwirth, "The Estimation of the Lorenz Curve and Gini Index," *Review of Economics and Statistics* 54 (1972): 306–316.

32. S. Albahra, T. Gorbett, S. Robertson, et al., "Artificial Intelligence and Machine Learning Overview in Pathology & Laboratory Medicine:

A General Review of Data Preprocessing and Basic Supervised Concepts," *Seminars in Diagnostic Pathology* 40, no. 2 (2023): 71–87.

33. J.-F. Lesevê, L. Bassegio, K. Maloum, et al., "Réalisation, Lecture et Interprétation Morphologique des Adénogrammes Et Empreintes Ganglionnaires. Enquête et Propositions du Groupe Francophone d'Hématologie Cellulaire (GFHC)," *Revue de Biologie Médicale* 357 (2020): 3–12.

34. T. Muzzafar, E. X. Wei, P. Lin, L. J. Medeiros, and J. L. Jorgensen, "Flow Cytometric Immunophenotyping of Anaplastic Large Cell Lymphoma," *Archives of Pathology & Laboratory Medicine* 133, no. 1 (2009): 49–56.

35. S. Schmid, M. Tinguely, P. Cione, H. Moch, and B. Bode, "Flow Cytometry as an Accurate Tool to Complement Fine Needle Aspiration Cytology in the Diagnosis of Low Grade Malignant Lymphomas," *Cytopathology* 22, no. 6 (2011): 397–406.

36. H. C. Bertram, I. J. Check, and M. A. Milano, "Immunophenotyping Large B-Cell Lymphomas: Flow Cytometric Pitfalls and Pathologic Correlation," *American Journal of Clinical Pathology* 116, no. 2 (2001): 191–203.

37. H. Fang and S. A. Wang, "Flow Cytometry Applications in the Diagnosis and Classification of Aggressive B-Cell Lymphoma," in *Precision Molecular Pathology of Aggressive B-Cell Lymphomas* (Springer, 2024), 125–142.

38. H. Matsushita, N. Nakamura, Y. Tanaka, et al., "Clinical and Pathological Features of B-Cell Non-Hodgkin Lymphomas Lacking the Surface Expression of Immunoglobulin Light Chains," *Clinical Chemistry and Laboratory Medicine* 50, no. 9 (2012): 1665–1670.

39. M. V. Kesler, G. S. Paranjape, S. L. Asplund, R. W. McKenna, S. Jamal, and S. H. Kroft, "Anaplastic Large Cell Lymphoma: A Flow Cytometric Analysis of 29 Cases," *American Journal of Clinical Pathology* 128, no. 2 (2007): 314–322.

40. M. R. Wilson, A. Barrett, C. Y. Cheah, and T. A. Eyre, "How I Manage Mantle Cell Lymphoma: Indolent Versus Aggressive Disease," *British Journal of Haematology* 201, no. 2 (2023): 185–198.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.