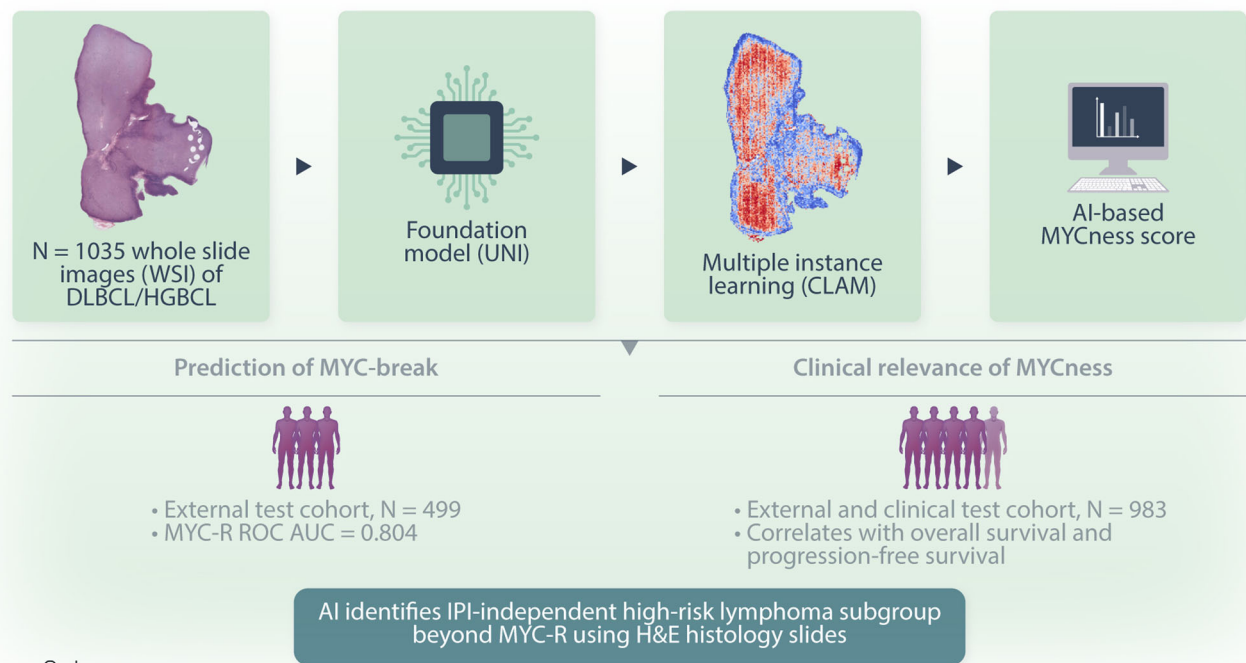


AI for prognostic assessment of diffuse large B-cell lymphoma using H&E whole-slide images

Jonas Lippl^{1,^}  | Sarah Reinke^{2,^} | Stefan Schrod³ | Lukas Wolfseher⁴ | Paul Hüttl⁴ | Michael Huttner⁴ | Annette M. Staiger^{5,6,7} | Ilske Oschlies² | Karoline Koch² | Edward G. Michaelis^{2,8} | Julia Richter² | Hilka Rauert-Wunderlich⁹ | Nicole Seifert¹ | Robin Kosch^{10,11} | Norbert Schmitz¹² | Marita Ziepert¹³  | Lorenz Thurner¹⁴  | Viola Poeschel¹⁴ | Bertram Glaß¹⁵ | Lorenz Trümper¹⁶ | Gerhard Held¹⁷ | David W. Scott¹⁸ | Andreas Rosenwald⁹ | German Ott⁵ | Rainer Spang⁴ | Michael Altenbuchinger^{1,^^} | Wolfram Klapper^{2,^^} 

Graphical Abstract

AI-derived morphologic MYCness predicts outcome in DLBCL/HGBCL



AI for prognostic assessment of diffuse large B-cell lymphoma using H&E whole-slide images

Jonas Lippl^{1,^}  | Sarah Reinke^{2,^} | Stefan Schrod³ | Lukas Wolfseher⁴ | Paul Hüttl⁴ | Michael Huttner⁴ | Annette M. Staiger^{5,6,7} | Ilske Oeschlies² | Karoline Koch² | Edward G. Michaelis^{2,8} | Julia Richter² | Hilka Rauert-Wunderlich⁹ | Nicole Seifert¹ | Robin Kosch^{10,11} | Norbert Schmitz¹² | Marita Ziepert¹³  | Lorenz Thurner¹⁴  | Viola Poeschel¹⁴ | Bertram Glaß¹⁵ | Lorenz Trümper¹⁶ | Gerhard Held¹⁷ | David W. Scott¹⁸ | Andreas Rosenwald⁹ | German Ott⁵ | Rainer Spang⁴ | Michael Altenbuchinger^{1,^^} | Wolfram Klapper^{2,^^} 

Correspondence: Wolfram Klapper (wkklapper@path.uni-kiel.de)

Abstract

Prediction of the outcome of large B-cell lymphomas/high-grade B-cell lymphomas (DLBCLs/HGBCLs) is based on clinical parameters and molecular testing, for example, for rearrangements of MYC (MYC-R) and MYC-R in combination with BCL2 and BCL6 translocations (double/triple hit). However, the group of DLBCL/HGBCL with poor outcome is not confined to MYC-R lymphomas, and fluorescence in situ hybridization (FISH) testing for MYC-R misses several high-risk lymphomas. We aimed to understand if artificial intelligence (AI) trained to identify MYC-R will delineate a poor prognostic subgroup. We generated a collection of digital hematoxylin and eosin (H&E)-stained slides of DLBCL/HGBCL ($N = 2018$) annotated for MYC, BCL2, and BCL6 translocations. A multiple-instance deep learning AI model for the identification of MYC-R alone or as double/triple hit was established using 1035 H&E-stained slides and evaluated on an external test cohort ($N = 499$). A pretrained tumor tissue classifier improved reliability and interpretability by focusing the model on tumor areas. Our model score reflects a morphological “MYCness” in DLBCL/HGBCL and demonstrates a strong association with overall survival (OS) and progression-free survival (PFS) in the external test cohort. This was confirmed with an additional clinical test cohort ($N = 484$) without FISH labels. The AI model scores correlate with various molecular features of DLBCL/HGBCL, including BCL2 and MYC gene expression, and the high-grade gene expression signature, but also features of the tumor microenvironment. Multivariate analysis, adjusted for International Prognostic Index (IPI) factors, demonstrated the prognostic significance of our model in identifying high-risk cases for both PFS and OS.

¹Department of Medical Bioinformatics, University Medical Center Göttingen, Göttingen, Germany

²Department of Pathology, Hematopathology Section, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany

³Department of Statistical Genomics and Systems Genetics, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany

⁴Department of Statistical Bioinformatics, University of Regensburg, Regensburg, Germany

⁵Department of Pathology, Robert Bosch Hospital Stuttgart, Stuttgart, Germany

⁶Dr Margarete Fischer-Bosch Institute of Clinical Pharmacology, Stuttgart, Germany

⁷University of Tuebingen, Tübingen, Germany

⁸Department of Pathology, Charité University Hospital Berlin, Berlin, Germany

⁹Institute of Pathology, University of Würzburg, Würzburg, Germany

¹⁰Peter L. Reichertz Institute for Medical Informatics, TU Braunschweig, Braunschweig, Germany

¹¹Hannover Medical School, Hannover, Germany

¹²Medical Clinic A, University Hospital Münster, Münster, Germany

¹³Department of Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany

¹⁴Department of Internal Medicine I, University Hospital of Saarland, Homburg, Germany

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *HemaSphere* published by John Wiley & Sons Ltd on behalf of European Hematology Association.

INTRODUCTION

A subgroup of large B-cell lymphomas (DLBCLs) harbors translocations involving the *MYC* gene (*MYC*-rearranged, *MYC*-R). These lymphomas are either classified as diffuse DLBCL or high-grade B-cell lymphoma (HGBCL) depending on the morphology and additional genetic features such as the co-occurrence of *BCL2* and *BCL6* translocations as double hit (DH) and triple hit (TH), respectively.^{1,2} The formal classification and nomenclature of DH and TH lymphomas differ slightly between the WHO¹ and the ICC² classifications. In this study, all DLBCLs are collectively referred to as DLBCL/HGBCL. Approximately 10% of lymphomas with the DLBCL/HGBCL morphology carry *MYC* breaks/translocation and are characterized by an unfavorable outcome.³ Consequently, screening of DLBCLs for *MYC* translocations by fluorescence in situ hybridization (FISH) has become a standard diagnostic procedure.^{1,2,4} Once a DLBCL/HGBCL turns out to be *MYC*-R, the specimen is usually subsequently tested for *BCL2* and *BCL6* translocations.⁵

However, single-gene molecular assays such as FISH to identify prognostically unfavorable DLBCL/HGBCL do not appear to have a promising future in DLBCL/HGBCL risk stratification for several reasons. First, molecular testing is not ubiquitously available in pathology laboratories, which may lead to delays in diagnosis and may also result in considerable costs. Thus, attempts have been undertaken to screen for cases by immunohistochemistry for *MYC* protein and restricting subsequent FISH analysis to a limited number of biopsies.^{6,7} Despite these efforts, it is now generally accepted that the presence of *MYC* translocations cannot be reliably detected by *MYC* immunohistochemistry under common diagnostic conditions.⁸ Second, and most importantly, the group of *MYC*-R/DH/TH hit lymphomas comprises only a subgroup of DLBCL/HGBCL with poor outcome, and FISH testing misses a relevant number of cases that fail under standard therapy.³ However, DH/TH DLBCL/HGBCL have successfully been used as a blueprint to identify lymphomas with a high risk of relapse using gene expression profiling for a “double-hit signature,” which also identifies lymphomas without *MYC*-R with high risk.^{9,10} Unfortunately, this technology is not widely available in most diagnostic labs.

A cost-efficient method is required to detect *MYC*-R as a surrogate for high-risk lymphomas in laboratories without expertise in molecular biology or access to the necessary technology, given the clinical need to identify the poor prognostic subgroup. One obvious approach to meeting this medical need is AI-based *MYC*-R detection from hematoxylin and eosin (H&E)-stained slides, which are generated for all patients by default during the diagnostic procedure. Deep learning methods are a key component of AI. The application of deep learning to whole-slide images (WSIs)—known as computational pathology—is a rapidly evolving field. Applications of computational pathology to H&E WSIs include tumor detection,^{11,12} survival analysis,^{13–17} or tumor subtyping.¹⁸ Recent approaches have overcome limitations such as a lack of annotations at the pixel (px) level¹⁹ or access to graphics processing units (GPUs).²⁰ Multiple-instance learning (MIL), a weakly supervised approach, utilizes the WSI label (diagnosis) to determine patches of the slide that are most relevant for the diagnosis instead. This approach has been the most successful for pathology image analysis.^{20–22}

However, despite this progress in digital pathology, success in detecting a single genetic aberration such as *MYC*-R in DLBCL/HGBCL analyzing H&E WSI is limited.^{23–27} Shortcomings of published studies include a high false-positive detection rate (AI predicting *MYC* in cases that do not harbor *MYC*-R) and a lack of clinical testing of these models.^{23–27}

In the current study, we investigated whether the false-positive rate of AI detecting *MYC* on H&E WSI can be reduced using the so far largest cohort of DLBCL/HGBCL H&E WSI and a novel filtering step, ensuring that AI models focus on lymphoma areas. However, our main goal was to understand whether the AI-based *MYC*-R prediction as a continuous score of morphological “*MYC*ness” could identify clinically aggressive disease, similar to the high-grade signature by gene expression profiling. Thus, we tested the AI score as a continuous parameter for its prognostic power in patients treated in prospective clinical trials.

MATERIALS AND METHODS

This study was designed in accordance with the standards for reporting clinical prediction models (Supporting Information S2: TRIPOD + AI Statement).

Whole-slide images and genetic annotations

The overall workflow of our study is illustrated in Figure 1. Three cohorts of H&E WSI pathologically reported as DLBCL or HGBCL specimens were generated to enable the development of a deep learning-based histomorphological score for the presence of *MYC*-R and/or DH/TH and to evaluate its association with clinical outcome (Supporting Information S1: Table 1). The first cohort, referred to as the development cohort, comprises *N* = 1035 DLBCL/HGBCL specimens, collected between 2000 and 2023 from the files of the Hematopathology Section in Kiel, Germany. Each case was reviewed by expert hematopathologists and tested for *MYC*-R using FISH by the break-apart assay, as previously described.²⁸ Lymphomas showing *MYC*-R were further analyzed for breaks in *BCL2* and *BCL6* to differentiate between *MYC*-R, DH, and TH lymphomas, and all categories served as molecular reference labels for model training.³ For the development cohort, no clinical follow-up was available. The second and third cohorts were established from patients enrolled in several prospective clinical trials of the German High Grade Lymphoma Study Group/German Lymphoma Alliance (GLA), including Denser²⁹ (*N* = 90), MINT³⁰ (*N* = 103), MegaCHOEP³¹ (*N* = 28), Ricover-60³² (*N* = 122), Ricover-noRTh³³ (*N* = 115), Smarter³⁴ (*N* = 141), and Unfolder³⁵ (*N* = 385). These two test cohorts consisted of a total of *N* = 983 patients treated with immunochemotherapy, including anti-CD20 antibodies. Follow-up information was available for these patients regarding overall survival (OS) and progression-free survival (PFS), as well as clinical data such as the International Prognostic Index (IPI), enabling an independent evaluation of the prognostic relevance of the model score. Within these two test cohorts, FISH annotations were available for *N* = 499 lymphomas^{36,37} (external test cohort). The remaining 484 cases with available WSIs and clinical data, but lacking genetic annotations, were referred to as the “Clinical

¹⁵Department of Hematology, Oncology, and Tumor Immunology, Helios Klinikum Berlin-Buch, Berlin, Germany

¹⁶Department of Hematology, Oncology, Georg August University of Göttingen, Göttingen, Germany

¹⁷Department of Internal Medicine, Westpfalz-Klinikum, Kaiserslautern, Germany

¹⁸BC Cancer and University of British Columbia, Vancouver, British Columbia, Canada

[^]Jonas Lippl and Sarah Reinke contributed equally.

^{^^}Michael Altenbuchinger and Wolfram Klapper shared senior authorship.

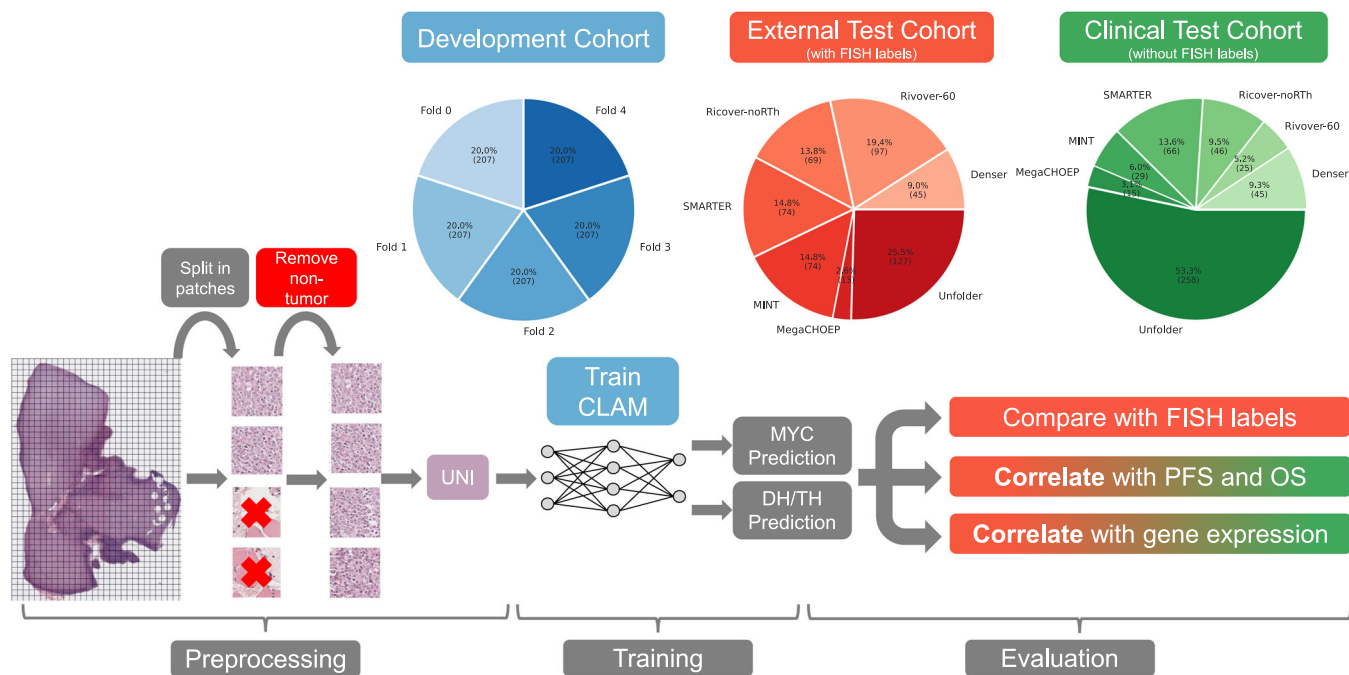


FIGURE 1 Workflow of this paper. Six multiple-instance learning (MIL) architectures are benchmarked, with CLAM providing the best performance. Predicted scores are subsequently evaluated on fluorescence in situ hybridization (FISH) labels of the external test cohort and correlated with clinical data of the external and clinical test cohort. Furthermore, correlations between gene expression values and the double-hit signature were calculated for a subgroup of patients of the external test cohort and the clinical test cohort with available data. DH, double hit; OS, overall survival; PFS, progression-free survival; TH, triple hit.

Test Cohort" (Figure 1). The number of WSIs was slightly higher than the number of patients mentioned above, because approximately 100 patients in the external test cohort had multiple images available. In these cases, predictions were averaged across images, and only one value per patient was used for further analysis. For study protocols, register numbers, and other details of the clinical trials setting, refer to their original publications.^{29–35} The development cohort was enriched for MYC-R and DH/TH cases, whereas the external and clinical test cohorts represent the prevalence of these genetic alterations in adults in Central Europe.³⁸

Data preprocessing

WSIs were converted from TIFF files into an SQLite database Pamly,³⁹ an open-source, in-house-developed library built on OpenSlide.⁴⁰ Only tiles containing tissue were stored, and white background areas were removed. Tiles measuring 512×512 pixels were sampled at the highest resolution, which corresponds to a magnification level of $400\times$ of a conventional light microscope. For each tile, a 1024-dimensional feature embedding was generated using the pretrained model known as UNI.⁴¹ UNI is a Vision Transformer⁴² that was pretrained in a self-supervised manner on 100,000 WSIs from various tissue types. The full data preprocessing pipeline can be found at <https://gitlab.gwdg.de/jlpl/cpathpipeline>. To facilitate tumor tissue classification, we curated a set of $N = 28,000$ patches that were manually annotated by expert pathologists. The tumor patch labels included "Tumor," "Tumor, Poor Quality," and "Partial Tumor" for tumor-like patches, while non-tumor patches were labeled as "Fatty Tissue," "Connective Tissue," and others (Supporting Information S1: [Supplementary Methods](#)). These labeled patches were used to train a tumor classifier, which was subsequently used to generate tumor probability maps and to pre-filter patches for downstream classification tasks.

Model development and validation

In total, we evaluated six different MIL methods on the training data using a fivefold cross-validation, to identify a model capable of learning robust histomorphological representations associated with aggressive lymphoma biology. The MYC-R and DH/TH status served as training labels. Tested methods were ABMIL,²¹ CLAM,²⁰ TransMIL,²² AttriMIL,⁴³ Snuffy,⁴⁴ and DTFD-MIL⁴⁵ (Supporting Information S1: [Supplementary Methods](#)). The best model was selected based on the area under the curve (AUC) of the receiver operating characteristic (ROC) in cross-validation using the training data and subsequently tested on the external test cohorts ($N = 499$). Due to the imbalance in the classes (Supporting Information S1: Table 1), balanced accuracy was used as a robust, unbiased metric to establish the thresholds for the dichotomized classification score. To determine if direct optimization for survival could further enhance prognostic performance, the best performing model was additionally fine-tuned on the external test cohort using DeepSurv⁴⁶ loss for both OS and PFS. The resulting fine-tuned model was subsequently evaluated on the independent clinical test cohort.

Clinical and molecular analyses

We assessed the correlation between the score derived from the selected model and OS or PFS by stratifying patients into two or four groups using the optimal threshold or quartiles, respectively. Kaplan-Meier (KM) curves were then used to compare OS and PFS probabilities among the subgroups. A log-rank test was used to calculate P-values and the concordance index (C-index) was used to quantify the correlation. The relationship between risk factors, trial, treatment, cell-of-origin (COO) subtype, and outcomes was determined using a Cox proportional hazards (CoxPH) model. Likelihood ratio tests (LRTs)

were used to evaluate the independent prognostic contribution of the model score. The correlation between molecular features such as gene expression and the model score was evaluated using Pearson's correlation coefficient.

RESULTS

Model development and primary validation

In total, six deep learning architectures were developed to learn histomorphological patterns associated with aggressive disease biology in DLBCL, using MYC-R and DH/TH status as training labels. A comprehensive benchmark analysis was performed for model selection using fivefold cross-validation on the development cohort of $N = 1035$ DLBCL/HGBCL (Figure 1 and Supporting Information S1: Table 1). CLAM²⁰ yielded the best performance, with a cross-validation ROC AUC (Supporting Information S1: Figure 1) of 0.824 ± 0.036 for MYC-R classification, and was selected as the final model (Supporting Information S1: Table 2). The corresponding performance evaluation in terms of area under the precision recall curve (AUPRC) supports this choice (Supporting Information S1: Table 2). For consistency, models for DH/TH prediction were established using the same workflow including a benchmark analysis (Supporting Information S1: Table 3). The latter confirmed CLAM to be one of the most suitable training strategies. Decision thresholds of 0.30 and 0.19 were finally established according to the maximal balanced accuracy of 0.75 (Supporting Information S1: Figure 2A) and 0.77 for MYC-R and DH/TH classification, respectively. A model that only utilized MYC/BCL2-DH lymphomas showed a similar performance (Supporting Information S1: Table 4).

Visualization of the models was performed by extraction of attention maps according to the procedure of Lu et al²⁰ (Figure 2). This visualization indicated that the model assigns high attention to areas of the WSI, which do not represent well-preserved lymphoma tissue according to a pathologist's visual evaluation (Figure 2A). To improve the plausibility of AI predictions, we restricted attention maps to regions identified by an expert pathologist-trained lymphoma pre-filtering algorithm. The resulting new visualization successfully focuses attention on regions likely containing tumor cells, while image artifacts, such as blurred patches and those at the boundary of the tissue specimen, are excluded (Figure 2B). It is worth noting that restricting the analysis to tiles containing lymphoma tissue and excluding irrelevant areas according to our pathologist-pretrained algorithm did not improve the model's ability to predict MYC-R when applied prior to the training (Supporting Information S1: Table 2). Nevertheless, the algorithm generates a result that enables pathologists to understand how the model works and likely increases confidence in the results.

External testing

The established models for MYC-R and DH/TH were subsequently tested using the external test cohort that included WSIs from $N = 499$ cases with corresponding genetic annotations (workflow Figure 1). Our AI models for MYC-R and DH/TH prediction were confirmed with an ROC AUC of 0.804 (Supporting Information S1: Figure 1 and Supporting Information S1: Table 2) and 0.778, respectively (Supporting Information S1: Table 3). The established decision thresholds yielded balanced accuracies of 0.75 (Supporting Information S1: Figure 2A) and 0.54, respectively. With these thresholds, our model predicted 41/44 (93%) MYC-R (true positive) and missed 3/44 (7%) cases with MYC-R by FISH as false negative. Notably, 319/455

(70%) of cases without MYC-R by FISH showed a high model score above the applied threshold, resulting in poor model calibration (Supporting Information S1: Figure 3). This finding suggests that the model score to which we refer as morphological "MYCness" is high in a subgroup of DLBCL/HGBCL without MYC-R/DH/TH.

Clinical relevance

To ensure that the model focused on tumor morphology, we used the tumor-filtered model for all subsequent analyses. Since our model score is the highest in lymphomas with MYC-R but also frequently increased in lymphomas lacking MYC-R, we reasoned that despite training for MYC-R, our model score reflects AI-detectable morphological features ("MYCness") shared by a group of lymphomas with and without MYC-R. In analogy to the gene expression high-grade signature, we hypothesized that a high score of morphological "MYCness" should be associated with poor outcome independent of the presence of MYC-R. Thus, we tested for an association of model scores with PFS and OS on the external test cohort via Cox regression, in which the respective model score serves as a continuous predictor variable (Supporting Information S1: Table 5). The most significant association was found between the MYC-R prediction score and PFS and OS ($P < 0.001$) even exceeding the prognostic value of the genetic status as MYC-R/DH/TH assessed by FISH. It is noteworthy that the MYC-R prediction score showed a stronger association with PFS and OS compared to the DH/TH score (Supporting Information S1: Table 7). To further optimize the model for predicting OS and PFS, we fine-tuned it by training the final layers with DeepSurv⁴⁶ loss on the external test cohort. However, the fine-tuned model to the independent clinical test cohort did not provide any additional prognostic information beyond the original MYC-R score. The C-index for OS and PFS was the same or slightly lower for the original and fine-tuned models. Therefore, we conclude that training an AI-based MYC-R score already achieved a prognostic performance that could not be further improved by fitting with the clinical outcome in the available data set. Therefore, the MYC-R score was used for all subsequent analyses.

To explore potential confounding variables, additional Cox regression analyses including international prognostic score (IPI) risk groups (low, low-intermediate, high-intermediate, and high), the trial, treatment, and COO subtype as additional predictor variables were conducted. In the analysis including only IPI as an additional variable, the model score retained significance (P -values 0.002 for OS and <0.001 for PFS). Interestingly, the prognostic relevance of the model score for PFS also remained significant when the respective ground-truth genetic annotation (the FISH result) was included in the Cox analysis as an independent variable (Supporting Information S1: Table 5, Supporting Information S1: Figure 4). Significance was also retained for the remaining analyses, including those for trial, treatment, and COO subtypes. This finding indicates that the trained score identifies histological features that are more closely linked to clinical aggressiveness than the objective of learning, MYC-R, itself. This is also confirmed by a stratification within the MYC- and MYC+ subgroups, in which the model can still identify significant low- and high-risk patients (Supporting Information S1: Figure 5). Furthermore, in another Cox regression analysis, we confirmed that the score is not confounded by the mere identification of trials or treatment groups. To verify the results on independent data, the model was applied to the clinical test cohort, which included patients from prospective clinical trials for which genetic annotations were not available. The Cox regression model shows P -values of 0.003 for OS and <0.001 PFS, confirming the prognostic value of the model score independent

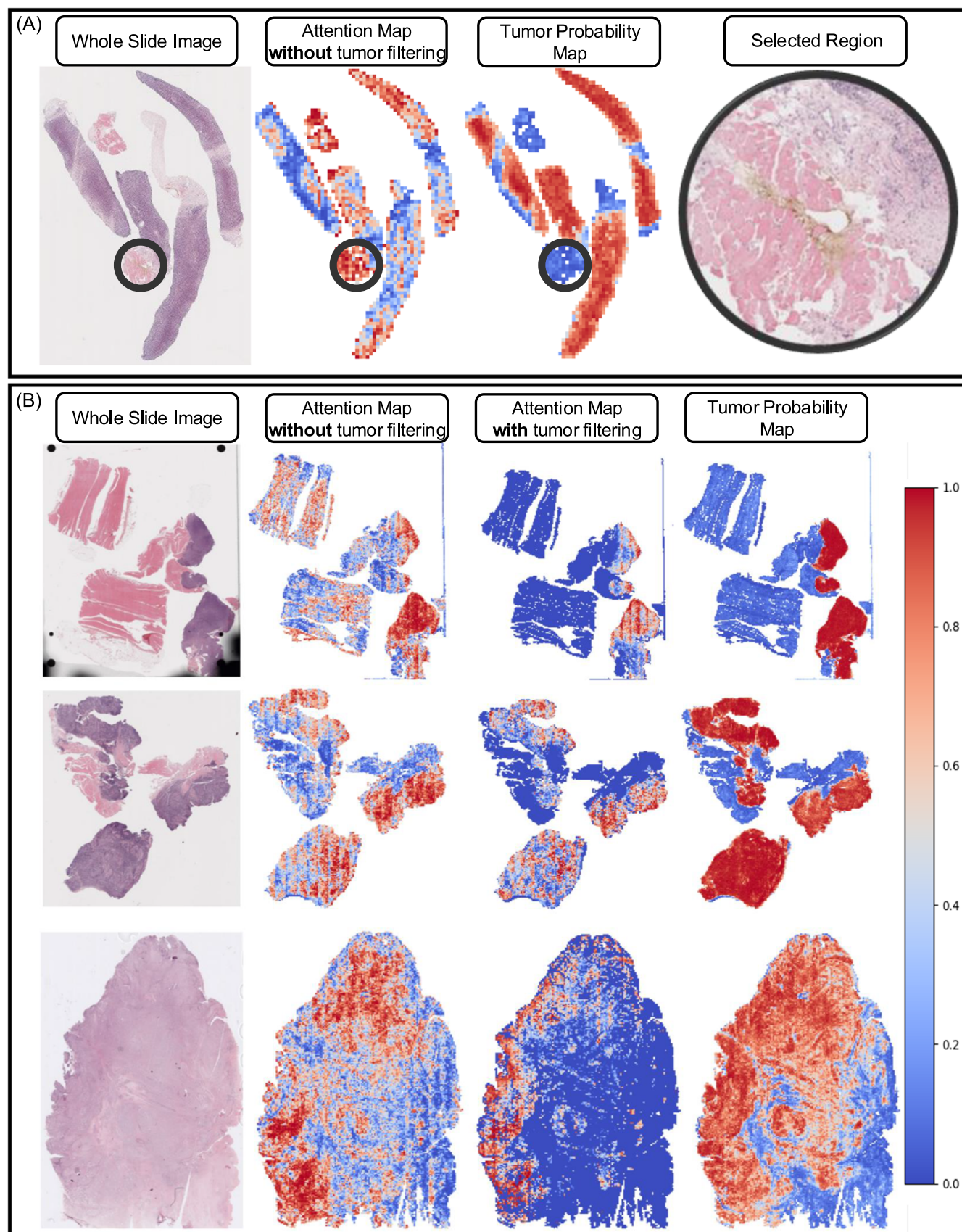


FIGURE 2 (See caption on next page).

FIGURE 2 Illustration of attention scores across different image regions. (A) Detailed illustration of non-tumor regions with unwanted high attention scores (visualized as percentiles per slide) by the model. The selected region in the hematoxylin and eosin (H&E)-stained whole-slide image (WSI) (black circle) receives high attention scores by the model, but has a very low tumor probability. High magnification reveals skeletal muscle and connective tissue, but no lymphoma. (B) The figure consists of three WSIs, each shown in a separate row: the original H&E-stained WSI, the model's attention map without tumor pre-filtering, the model's attention map after tumor pre-filtering, and the tumor probability map. The color scale on the right indicates the respective score, ranging from red (high) to blue (low). Without pre-filtering, high attention is assigned to eosinophilic areas, which represent smooth muscle tissue (upper row) or necrosis (middle row).

of MYC-R. For the remaining analyses, including those for treatment and trial, we observed significant P-values for PFS and OS. In contrast, if COO is included as an additional variable, the model score was rendered insignificant. This, however, might be attributable to the comparable low number of available COO labels ($N = 449$). To further evaluate the prognostic value of the model score, we used LRTs by fitting a full model that included the variable of interest and an adjusted reduced model that excluded that variable (Supporting Information S1: Table 6). Consistent with previous findings, the model score showed significant prognostic improvement, except for OS in the clinical test cohort. Our model score is a continuous variable and does not dichotomize patients like FISH for MYC-R/DH/TH. To analyze the score in a dichotomized fashion, the external test cohort and the clinical test cohort (for baseline characteristics, see Supporting Information S1: Table 8) were reconciled and systematically tested for an association between dichotomized model scores on the one hand and OS and PFS on the other hand. Significance levels were calculated using log-rank tests. The threshold yielding the strongest separation within the external test cohort between high- and low-risk groups was selected and subsequently applied to stratify the clinical test cohort into high- and low-risk groups. This optimization of patient stratification into risk groups identifies a high-risk group with a median PFS of 47.6 months (44.7% 5-year PFS) and a median PFS that has not been reached (77.3% 5-year PFS) in the low-risk group (Figure 3). Similarly, for OS, we observed a 5-year OS rate of 54.3% for the high-risk patients and 82.4% for the low-risk patients (Figure 4).

Molecular features associated with “MYCness”

Given the strong association between the model score and survival in lymphomas even without MYC-R, we aimed to evaluate the biological context of the score. All model scores (MYC-R or DH/TH) were significantly associated with the presence of the DH/dark zone signature (coef = 0.286, $P < 0.001$).⁴⁷ A positive association was detected with the ABC COO subtype (coef = 0.138, $P = 0.003$) and vice versa for the GCB subtype. For 509 cases in the external and clinical cohorts, DHIT Nanostring gene expression data for 90 genes were available (Supporting Information S1: Table 9). For data pre-processing, a pseudo count was added and profiles were normalized to an average count of 1000. Finally, data were $\log_2$ -transformed. In addition, preprocessed gene expression for the lymphoma-associated macrophage interaction signature (LAMIS)⁴⁸ for an additional 269 cases was used (Supporting Information S1: Table 10). A positive correlation was found between model scores and MYC and BCL2. Moreover, the model score was inversely correlated with the expression of genes reflecting macrophages, for example, CD274, CD14, and CD68, in line with our previous data demonstrating that a high macrophage content in the tumor microenvironment is associated with a favorable outcome.⁴⁹ Moreover, genes reflecting the microenvironment, such as the CD30 ligand (TNFSF8), were also associated with the model score. Thus, despite the fact that our score was trained to identify cases with MYC-R, it reflects multiple

biological features of neoplastic cells, including MYC expression, the DH signature, and the tumor microenvironment.

DISCUSSION

DLBCL/HGBCL are the most common group of lymphomas in adults. Subtypes of DLBCL/HGBCL such as the COO subtypes or MYC-R require molecular testing to be identified. Clinically relevant subtypes are DLBCL/HGBCL with MYC-R, especially DH/TH, as they show an unfavorable outcome under standard immunochemotherapy³ and are thus recognized either as distinct prognostic subgroups or as entities by current classifications.^{1,4} Even specialized pathologists are unable to recognize MYC-R lymphoma using H&E slides.^{2,50} Auxiliary techniques such as immunohistochemistry for MYC protein were not helpful under diagnostic conditions.⁵¹ Thus, current diagnostic guidelines recommend screening all lymphomas with the morphology of DLBCL/HGBCL for MYC-R by FISH or other suitable molecular tests.⁵ However, the analysis of MYC-R/DH/TH by FISH has several disadvantages in risk stratification of patients with DLBCL/HGBCL. First, FISH may miss a subset of lymphomas with MYC-R.^{9,52} Second, although FISH is a standard diagnostic procedure in many laboratories, access may be limited in low- and middle-income countries.^{53,54} Third, and most importantly, molecular detection of MYC-R is only demarcating a small subgroup of <10% of all patients with DLBCL/HGBCL, and a considerable number of patients with high risk of relapse and unfavorable outcome remain undetected within the group of MYC-negative lymphomas.^{55,56} To further dissect MYC-negative lymphomas into prognostic subgroups, gene expression profiling or next-generation sequencing is required.^{47,55} Here, we demonstrate that application of AI to H&E-stained slides may help to overcome some of the aforementioned limitations of molecular profiling at least with respect to risk assessment.

We first asked whether the prediction of MYC-R by AI using H&E WSI can be improved compared to previous studies^{23–27} to a level allowing substitution of molecular testing. However, although we overcame several of the technical limitations of previous studies by generating the largest data set of WSI of DLBCL/HGBCL ($n > 2000$) currently available worldwide, testing multiple state-of-the-art MIL methods and automated restriction to lymphoma-containing areas of the WSI, AI-predicted MYC-R/DH/TH still suffers from moderate to low specificity, as many lymphomas without MYC-R/DH/TH receive high model scores of “MYCness.” Utilizing the model to select for specimen requiring FISH testing to confirm MYC-R/DH/TH in a diagnostic setting would reduce the number of tests to be conducted by only about 34% compared to the current standard approach of testing all DLBCL/HGBCL.

Our second aim was to understand if AI-based MYC prediction is capturing a common biology and clinical behavior in all DLBCL/HGBCL independent of MYC-R, as has been observed in gene expression profiling studies.⁹ Previous studies indicated that MYC-R/DH/TH DLBCL/HGBCL share gene expression features with a group of lymphomas lacking these genetic aberrations but showing the same poor prognosis.⁹ Assuming that a common gene expression

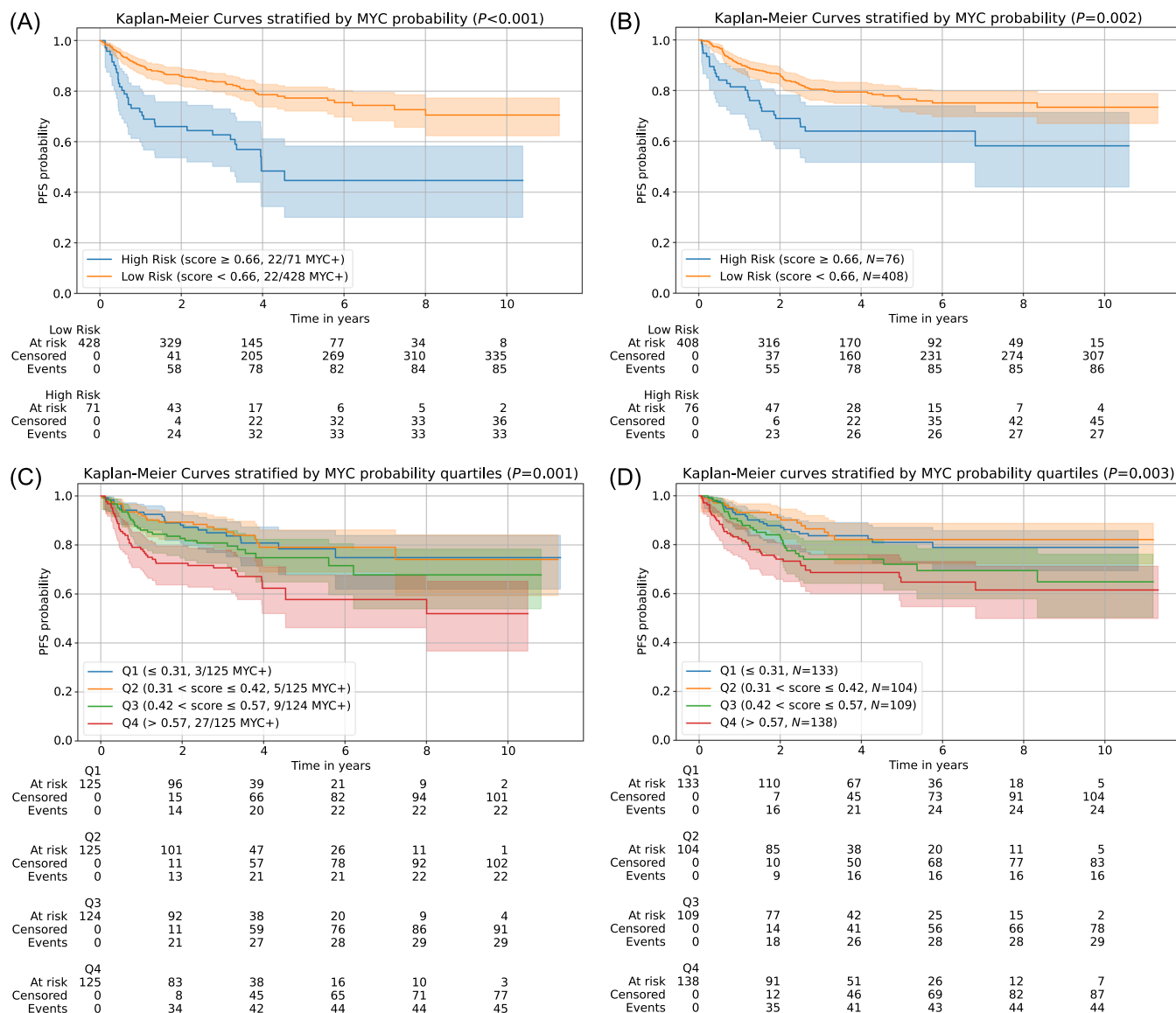


FIGURE 3 Kaplan-Meier curves for progression-free survival. Kaplan-Meier curves for progression-free survival (PFS) of the external test cohort (A, C) and the clinical test cohort (B, D) for patient stratification based on model scores. The numbers of MYC-R cases and total patients in each subgroup are listed in the legends. First, the threshold for stratification was selected using the minimum P-value on the external test cohort (A) and applied to the clinical test cohort (B). In addition, patients are stratified using quartiles of the model scores (C, D).

profile will be reflected in lymphoma morphology assessed by AI in the H&E WSI, we reasoned that our AI-based score will associate with outcome not despite but precisely because it identifies morphological features associated with MYC/DH/TH biology ("MYCness") in lymphomas lacking genetic MYC-R. It is noteworthy that AI-based "MYCness" does refer to morphological features detected and quantified by our AI approach, but not necessarily to biological MYC signaling. In line with our hypothesis, the AI-based model is in fact a better prognostic marker than the template MYC-R that it was trained on. We conclude that our AI-based "MYCness" score is a stronger predictor of outcome than the MYC status on which it was trained. The model retained prognostic significance independently of established clinical risk factors, such as the IPI, and identified high-risk patients with genetically MYC-R/DH/TH-negative lymphomas. This indicates that the model detects aggressive disease biology rather than genetic status alone.

The features selected by the model to identify "MYCness" remain elusive as an inherent feature of AI-based methods. However, the attention of the model to individual tiles of the WSI can be quantified and illustrated in attention maps. Visual inspection of attention maps indicates that most models that we applied assigned high attention to areas not considered to contain lymphoma tissue. Although necrosis may not represent lymphoma morphology, it may nevertheless be used by pathologists to identify lymphoma subtypes.⁵⁷ However, areas with improper staining or non-infiltrated tissue received high attention too. Thus, we deployed the expertise of pathologists and trained a tumor classifier that automatically excluded non-lymphoma tissue in the training process. This procedure also ensures that the model learns from tissue regions that pathologists consider relevant for diagnosis and reduces potential shortcomings due to learning from artifacts. As a result, the attention maps become histologically plausible by utilizing regions that the pathologists themselves would

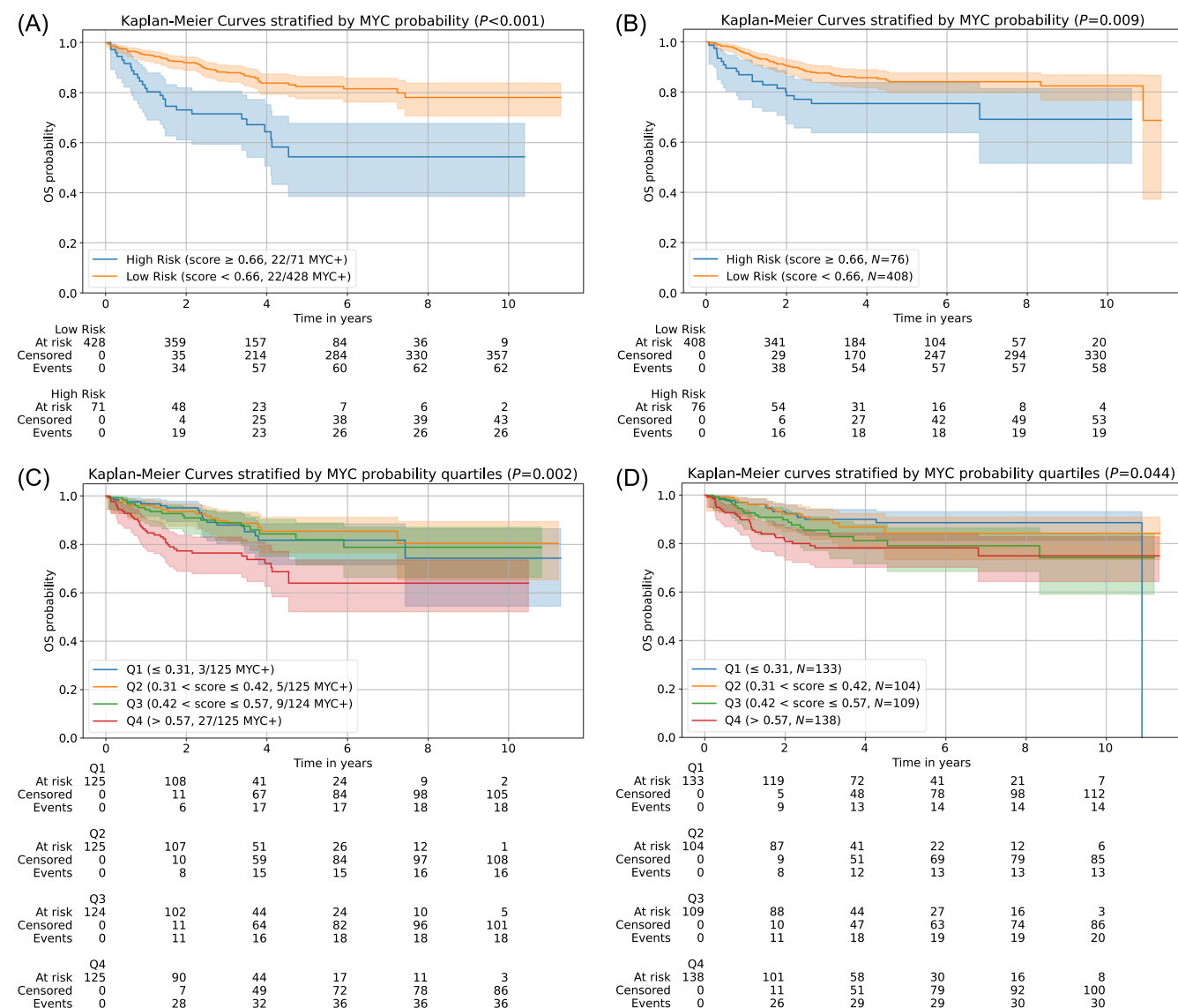


FIGURE 4 Kaplan-Meier curves for overall survival. Kaplan-Meier curves for overall survival (OS) of the external test cohort (A, C) and the clinical test cohort (B, D). Risk groups were defined based on model scores, using a cutoff selected via minimum P-value analysis of the external test cohort (A), and then applied to the clinical test cohort (B). Quartiles (C, D) were used for additional visualization and patient stratification. The numbers of MYC-R cases and patients in each subgroup are provided in the legends.

focus on in visual inspection. Although this workflow, surprisingly, does not improve the performance of the model, we believe that it improves the comprehensibility of the result and increases acceptance among pathologists, which will be crucial for future clinical applications.

As discussed above, the fact that our model was trained to identify MYC-R but also provides prognostic information in lymphomas lacking MYC-R does not essentially come as a surprise. In fact, the gene expression-based DH signature was developed similarly.⁹ DLBCL/HGBCL with MYC-R/DH/TH obviously show features in their H&E appearance that are associated with an aggressive clinical course independent of the molecular genetic background,⁵⁵ and these features seem to be captured by our model. Being trained to detect MYC-R/DH/TH, the association with the DH/dark zone gene expression signature^{9,10} as well as protein expression of MYC and BCL2^{58,59} with the “MYCness” seems plausible. It is important to

stress that “MYCness” in our study reflects a morphology-based score generated by AI on H&E slides, but not a score for biological activity of MYC, although both might be related. It is noteworthy that the model result is also associated with microenvironmental signatures known to provide molecular genetic subtype-overarching prognostic information, which have not yet been linked to MYC-R.^{49,60} Future studies will be required to identify the molecular features associated best with the model score and to understand if the training is reflecting more than just “MYCness” in the narrow sense. These future studies may also help to understand why training for MYC-R provided a better outcome predictor compared to training for DH/TH, although the latter were the core cohort to define the DH signature by gene expression profiling and are considered a molecularly rather homogeneous entity.^{9,61}

Training to detect MYC-R may appear as an indirect way to generate an outcome predictor by AI, raising the question of why AI

was not trained directly to predict clinical outcome. However, robust outcome-trained models from WSI require not only sufficiently large cohorts with clinical follow-up but also adequate numbers of outcome events, endpoint quality, cohort diversity, treatment heterogeneity, and adequate follow-up duration. Although our cohort of almost 1000 clinically annotated DLBCL/HGBCL cases is substantial, we considered it insufficient for reliable independent training and validation of a dedicated outcome-prediction model within the scope of the present study. Therefore, we took advantage of the fact that large cohorts of DLBCL/HGBCL with MYC-R/DH/TH are available from diagnostic testing and that the association of these features with outcome is well established. Nevertheless, joining WSI of DLBCL/HGBCL from multiple groups and consortia in the future is certainly desirable and may enable researchers to train AI directly for outcome prediction. So far, it remains uncertain whether such an approach will provide a better model than the one based on MYC presented in the current study. A limitation of our model as well as that of most published studies of this kind is a lack of knowledge on slide/tissue/scan features on the performance of the model. Recent studies suggest that foundation models are still susceptible to batch effects despite having diverse training data.⁶² However, the biopsies in our study were retrieved from pathology centers all over Germany, reflect the diagnostic conditions in this area, and demonstrate that the model is successful despite the heterogeneity of the tissue processing. Nevertheless, future studies should evaluate systematically the influence of tissue processing, cutting, staining, and scanning on model performance. In addition, potential benefits of color normalization and stain augmentation should be studied. The presented study showed that predictions were average when multiple slides per patient were available. Performance could be further improved with alternative aggregation strategies, such as focusing on high-quality slides assessed by pathologists.

Our model is certainly useful in a retrospective analysis to study risk profiles and clinical outcome over multiple independent cohorts without the requirement of additional molecular testing. Once the model's performance is confirmed in additional independent cohorts, it can be used in prospective clinical trials. The "MYCness" model may be clinically useful when targeting MYC⁶³ or when considering intensification of therapy for high-risk patients.⁶⁴ These potential use cases require the definition of cutoffs based on the purpose of the risk stratification.

In summary, we present the first morphology-based AI model that captures "MYCness" as an aggressive phenotype of DLBCL/HGBCL from routine H&E-stained WSIs. The predicted score provides prognostic information independent of the IPI, enabling large-scale risk stratification in prospective clinical trials for pathologists and researchers worldwide without the need for expert knowledge in the field.

ACKNOWLEDGMENTS

The authors would like to thank all the patients who took part in the Denser, Ricover-60, Ricover-noRTh, MINT, Smarter, Unfolder, and MegaCHOEP trials, as well as all the clinicians and pathologists who supported the trials.

AUTHOR CONTRIBUTIONS

Jonas Lippl: Conceptualization; writing—original draft; formal analysis; methodology; visualization; writing—review and editing; software. **Sarah Reinke:** Conceptualization; writing—review and editing. **Stefan Schrod:** Formal analysis; writing—review and editing. **Lukas Wolfseher:** Formal analysis; methodology; writing—review and editing. **Paul Hüttl:** Formal analysis; methodology; writing—review and

editing. **Michael Huttner:** Formal analysis; methodology; writing—review and editing. **Annette M. Staiger:** Writing—review and editing; data curation. **Ilse Oschlies:** Writing—review and editing; data curation. **Karoline Koch:** Writing—review and editing; data curation. **Edward G. Michaelis:** Writing—review and editing; data curation. **Julia Richter:** Writing—review and editing; data curation. **Hilka Rauert-Wunderlich:** Writing—review and editing; data curation. **Nicole Seifert:** Formal analysis; methodology; writing—review and editing. **Robin Kosch:** Formal analysis; methodology; writing—review and editing. **Norbert Schmitz:** Writing—review and editing; data curation. **Marita Ziepert:** Writing—review and editing; data curation. **Lorenz Thurner:** Writing—review and editing; data curation. **Viola Poeschel:** Writing—review and editing; data curation. **Bertram Glaß:** Writing—review and editing; data curation. **Lorenz Trümper:** Writing—review and editing; data curation. **Gerhard Held:** Writing—review and editing; data curation. **David W. Scott:** Writing—review and editing; data curation. **Andreas Rosenwald:** Writing—review and editing; data curation; funding acquisition. **German Ott:** Writing—review and editing; data curation; funding acquisition. **Rainer Spang:** Writing—review and editing; conceptualization; funding acquisition. **Michael Altenbuchinger:** Writing—review and editing; conceptualization; writing—original draft; supervision; funding acquisition. **Wolfram Klapper:** Conceptualization; writing—review and editing; writing—original draft; data curation; funding acquisition.

CONFLICT OF INTEREST STATEMENT

H.R.-W.: Speaker for Incyte Corporation. A.R.: Institutional research contract—MorphoSys and Incyte. S.R.: Institutional research funding from Incyte. W.K.: Institutional research funding from Incyte, Amgen, Roche, and Janssen. The other authors report no competing interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ETHICS STATEMENT

This study was conducted according to the recommendations of the Institutional Review Board of the Medical Faculty, University of Kiel (number D447/10), and is covered by the patient-informed consent of the respective trials included.

FUNDING

This work was supported by the German Federal Ministry of Education and Research (BMBF) as project FDLP under grant numbers 01KD2209A, 01KD2209B, 01KD2209C, 01KD2209D, and 01KD2209E, and project FDLP2 under grant numbers 01KD2415A, 01KD2415B, 01KD2415C, 01KD2415D, and 01KD2415E. Additional support was obtained from the Lunenburg Lymphoma Biomarker Consortium (LLBC). G.O. was additionally supported by the Robert-Bosch-Foundation, Stuttgart. Open Access funding enabled and organized by Projekt DEAL.

SUPPORTING INFORMATION

Additional supporting information can be found in the online version of this article.

ORCID

Jonas Lippl  <https://orcid.org/0009-0009-9492-3255>

Marita Ziepert  <https://orcid.org/0000-0002-0834-8111>

Lorenz Thurner  <https://orcid.org/0000-0002-8167-2989>

Wolfram Klapper  <https://orcid.org/0000-0001-7208-4117>

REFERENCES

- Alaggio R, Amador C, Anagnostopoulos I, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. *Leukemia*. 2022;36:1720-1748.
- Campo E, Jaffe ES, Cook JR, et al. The International Consensus Classification of Mature Lymphoid Neoplasms: a report from the Clinical Advisory Committee. *Blood*. 2022;140:1229-1253.
- Rosenwald A, Bens S, Advani R, et al. 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.
- Haematolymphoid Tumours: Who Classification of Tumours. International Agency for Research on Cancer; 2024.
- de Leval L, Alizadeh AA, Bergsagel PL, et al. Genomic profiling for clinical decision making in lymphoid neoplasms. *Blood*. 2022;140:2193-2227.
- Green TM, Nielsen O, de Stricker K, Xu-Monette ZY, Young KH, Möller MB. High levels of nuclear MYC protein predict the presence of MYC rearrangement in diffuse large B-cell lymphoma. *Am J Surg Pathol*. 2012;36:612-619.
- Chisholm KM, Bangs CD, Bacchi CE, Kirsch HM, Cherry A, Natkunam Y. Expression profiles of MYC protein and MYC gene rearrangement in lymphomas. *Am J Surg Pathol*. 2015;39:294-303.
- Obermann EC, Csato M, Dirnhofer S, Tzankov A. Aberrations of the MYC gene in unselected cases of diffuse large B-cell lymphoma are rare and unpredictable by morphological or immunohistochemical assessment. *J Clin Pathol*. 2009;62:754-756.
- Ennishi D, Jiang A, Boyle M, et al. Double-hit gene expression signature defines a distinct subgroup of germinal center B-cell-like diffuse large B-cell lymphoma. *J Clin Oncol*. 2019;37:190-201.
- Sha C, Barrans S, Cucco F, et al. Molecular high-grade B-cell lymphoma: defining a poor-risk group that requires different approaches to therapy. *J Clin Oncol*. 2019;37:202-212.
- Bandi P, Geessink O, Manson Q, et al. From detection of individual metastases to classification of lymph node status at the patient level: the camelyon17 challenge. *IEEE Trans Med Imaging*. 2019;38:550-560.
- Litjens G, Bandi P, Ehteshami Bejnordi B, et al. 1399 H&E-stained sentinel lymph node sections of breast cancer patients: the CAMLYON dataset. *Gigascience*. 2018;7:giy065.
- Zhu X, Yao J, Zhu F, Huang J. Wsisa: making survival prediction from whole slide histopathological images. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017:7234-7242.
- Tang B, Li A, Li B, Wang M. CapSurv: capsule network for survival analysis with whole slide pathological images. *IEEE Access*. 2019;7:26022-26030.
- Yao J, Zhu X, Jonnagaddala J, Hawkins N, Huang J. Whole slide images based cancer survival prediction using attention guided deep multiple instance learning networks. *Med Image Anal*. 2020;65:101789.
- Chen RJ, Lu MY, Shaban M, et al. Whole slide images are 2d point clouds: context-aware survival prediction using patch-based graph convolutional networks. In: *Medical Image Computing and Computer Assisted Intervention-MICCAI 2021: 24th International Conference, Strasbourg, France, September 27-October 1, 2021, Proceedings, Part VIII* 24, 2021:339-349.
- Ammeling J, Schmidt LH, Ganz J, et al. Attention-based multiple instance learning for survival prediction on lung cancer tissue microarrays. In: *BVM Workshop*, 2023:220-225.
- Hashimoto N, Fukushima D, Koga R, et al. Multi-scale domain-adversarial multiple-instance CNN for cancer subtype classification with unannotated histopathological images. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020:3852-3861.
- Campanella G, Hanna MG, Geneslaw L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med*. 2019;25:1301-1309.
- Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng*. 2021;5:555-570.
- Ilse M, Tomczak J, Welling M. Attention-based deep multiple instance learning. In: *International Conference on Machine Learning*, 2018:2127-2136.
- Qian Z, Li K, Lai M, et al. Transformer based multiple instance learning for weakly supervised histopathology image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2022:160-170.
- Syrykh C, Di Proietto V, Brion E, et al. MYC rearrangement prediction from LYSA whole slide images in large B-cell lymphoma: a multi-centric validation of self-supervised deep learning models. *Mod Pathol*. 2024;37(12):100610.
- Dooper S, Pinckaers H, Aswolinskiy W, et al. Gigapixel end-to-end training using streaming and attention. *Med Image Anal*. 2023;88:102881.
- Yüce A, Doerig C, Gildenblat J, et al. Identification of MYC-driven high-grade B-cell lymphoma using deep learning-based whole slide image analysis. *Blood*. 2023;142:3649.
- Perry C, Greenberg O, Haberman S, et al. Image-based deep learning detection of high-grade B-cell lymphomas directly from hematoxylin and eosin images. *Cancers*. 2023;15:5205.
- Swiderska-Chadaj Z, Hebeda KM, van den Brand M, Litjens G. Artificial intelligence to detect MYC translocation in slides of diffuse large B-cell lymphoma. *Virchows Arch*. 2021;479:617-621.
- Richter J, Hüttmann A, Rekowski J, et al. Molecular characteristics of diffuse large B-cell lymphoma in the Positron Emission Tomography-Guided Therapy of Aggressive Non-Hodgkin lymphomas (PETAL) trial: correlation with interim PET and outcome. *Blood Cancer J*. 2019;9:67.
- Murawski N, Pfreundschuh M, Zeynalova S, et al. Optimization of rituximab for the treatment of DLBCL (I): dose-dense rituximab in the DENSE-R-CHOP-14 trial of the DSHNHL. *Ann Oncol*. 2014;25:1800-1806.
- Pfreundschuh M, Trümper L, Österborg A, et al. CHOP-like chemotherapy plus rituximab versus CHOP-like chemotherapy alone in young patients with good-prognosis diffuse large-B-cell lymphoma: a randomised controlled trial by the MabThera International Trial (MInT) Group. *Lancet Oncol*. 2006;7:379-391.
- Schmitz N, Nickelsen M, Ziepert M, et al. Conventional chemotherapy (CHOEP-14) with rituximab or high-dose chemotherapy (MegaCHOEP) with rituximab for young, high-risk patients with aggressive B-cell lymphoma: an open-label, randomised, phase 3 trial (DSHNHL 2002-1). *Lancet Oncol*. 2012;13:1250-1259.
- Pfreundschuh M, Schubert J, Ziepert M, et al. Six versus eight cycles of bi-weekly CHOP-14 with or without rituximab in elderly patients with aggressive CD20+ B-cell lymphomas: a randomised controlled trial (RICOVER-60). *Lancet Oncol*. 2008;9:105-116.
- Held G, Murawski N, Ziepert M, et al. Role of radiotherapy to bulky disease in elderly patients with aggressive B-cell lymphoma. *J Clin Oncol*. 2014;32:1112-1118.
- Pfreundschuh M, Poeschel V, Zeynalova S, et al. Optimization of rituximab for the treatment of diffuse large B-cell lymphoma (II): extended rituximab exposure time in the SMARTE-R-CHOP-14 trial of the german high-grade non-Hodgkin lymphoma study group. *J Clin Oncol*. 2014;32:4127-4133.
- Thurner L, Ziepert M, Berdel C, et al. Radiation and dose-densification of R-CHOP in aggressive B-cell lymphoma with intermediate prognosis: the UNFOLDER study. *HemaSphere*. 2023;7:e904.

36. Staiger AM, Ziepert M, Horn H, et al. Clinical impact of the cell-of-origin classification and the MYC/BCL2 dual expresser status in diffuse large B-cell lymphoma treated within prospective clinical trials of the German high-grade non-Hodgkin's lymphoma study group. *J Clin Oncol*. 2017;35:2515-2526.
37. Horn H, Ziepert M, Wartenberg M, et al. Different biological risk factors in young poor-prognosis and elderly patients with diffuse large B-cell lymphoma. *Leukemia*. 2015;29:1564-1570.
38. Paul U, Richter J, Stuhlmann-Laietz C, et al. Advanced patient age at diagnosis of diffuse large B-cell lymphoma is associated with molecular characteristics including ABC-subtype and high expression of MYC. *Leuk Lymphoma*. 2018;59:1213-1221.
39. Huttner M. Pamly [Internet]. 2025. <https://pamly.spang-lab.de/>
40. Goode A, Gilbert B, Harkes J, Jukic D, Satyanarayanan M. OpenSlide: a vendor-neutral software foundation for digital pathology. *J Pathol Inform*. 2013;4:27.
41. Chen RJ, Ding T, Lu MY, et al. Towards a general-purpose foundation model for computational pathology. *Nat Med*. 2024;30:850-862.
42. Dosovitskiy A. An image is worth 16 × 16 words: transformers for image recognition at scale. *ArXiv*. Preprint 2020. ArXiv201011929.
43. Cai L, Huang S, Zhang Y, Lu J, Zhang Y. Rethinking attention-based multiple instance learning for whole-slide pathological image classification: an instance attribute viewpoint. *ArXiv*. Preprint. 2024. ArXiv240400351.
44. Jafarinia H, Alipanah A, Razavi S, Mirzaie N, Rohban MH. Snuffy: efficient whole slide image classifier. In: *European Conference on Computer Vision*, 2024:243-260.
45. Zhang H, Meng Y, Zhao Y, et al. Dtf-d-mil: double-tier feature distillation multiple instance learning for histopathology whole slide image classification. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022:18802-18812.
46. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med Res Methodol*. 2018;18:24.
47. Alduaij W, Collinge B, Ben-Neriah S, et al. Molecular determinants of clinical outcomes in a real-world diffuse large B-cell lymphoma population. *Blood*. 2023;141:2493-2507.
48. Staiger AM, Altenbuchinger M, Ziepert M, et al. A novel lymphoma-associated macrophage interaction signature (LAMIS) provides robust risk prognostication in diffuse large B-cell lymphoma clinical trial cohorts of the DSHNHL. *Leukemia*. 2020;34:543-552.
49. Croci GA, Au-Yeung RKH, Reinke S, et al. SPARC-positive macrophages are the superior prognostic factor in the microenvironment of diffuse large B-cell lymphoma and independent of MYC rearrangement and double-/triple-hit status. *Ann Oncol*. 2021;32:1400-1409.
50. Natkunam Y, de Jong D, Farinha P, et al. Lack of reproducibility of histopathological features in MYC-rearranged large B cell lymphoma using digital whole slide images: a study from the Lunenburg Lymphoma Biomarker Consortium. *Histopathology*. 2023;82:1105-1111.
51. Kluk MJ, Ho C, Yu H, et al. MYC immunohistochemistry to identify MYC-driven B-cell lymphomas in clinical practice. *Am J Clin Pathol*. 2016;145:166-179.
52. Chong LC, Ben-Neriah S, Slack GW, et al. High-resolution architecture and partner genes of MYC rearrangements in lymphoma with DLBCL morphology. *Blood Adv*. 2018;2:2755-2765.
53. Radich JP, Briercheck E, Chiu DT, et al. Precision medicine in low- and middle-income countries. *Annu Rev Pathol: Mech Dis*. 2022;17:387-402.
54. Valvert F, Silva O, Solórzano-Ortiz E, et al. Low-cost transcriptional diagnostic to accurately categorize lymphomas in low- and middle-income countries. *Blood Adv*. 2021;5:2447-2455.
55. Hilton LK, Scott DW, Morin RD. Biological heterogeneity in diffuse large B-cell lymphoma. *Sem Hematol*. 2023;60:267-276.
56. Collinge B, Ben-Neriah S, Chong L, et al. The impact of MYC and BCL2 structural variants in tumors of DLBCL morphology and mechanisms of false-negative MYC IHC. *Blood*. 2021;137:2196-2208.
57. Asano N, Yamamoto K, Tamaru JI, et al. Age-related Epstein-Barr virus (EBV)-associated B-cell lymphoproliferative disorders: comparison with EBV-positive classic Hodgkin lymphoma in elderly patients. *Blood*. 2009;113:2629-2636.
58. Bogusz AM, Kovach AE, Le LP, Feng D, Baxter RHG, Sohani AR. Diffuse large B-cell lymphoma with concurrent high MYC and BCL2 expression shows evidence of active B-cell receptor signaling by quantitative immunofluorescence. *PLoS One*. 2017;12:e0172364.
59. Johnson NA, Slack GW, Savage KJ, et al. 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.
60. Steen CB, Luca BA, Esfahani MS, et al. The landscape of tumor cell states and ecosystems in diffuse large B cell lymphoma. *Cancer Cell*. 2021;39:1422-1437.e10.
61. Scott DW, King RL, Staiger AM, et al. High-grade B-cell lymphoma with MYC and BCL2 and/or BCL6 rearrangements with diffuse large B-cell lymphoma morphology. *Blood*. 2018;131:2060-2064.
62. Kömen J, Marienwald H, Dippel J, Hense J. Do histopathological foundation models eliminate batch effects? A comparative study. Preprint 2024. doi:10.48550/arXiv.2411.05489
63. Stokes ME, Wenzl K, Huang CC, et al. Transcriptomic classification of diffuse large B-cell lymphoma identifies a high-risk activated B-cell-like subpopulation with targetable MYC dysregulation. *Nat Commun*. 2024;15:6790.
64. Dührsen U, Müller S, Hertenstein B, et al. Positron emission tomography-guided therapy of aggressive non-Hodgkin lymphomas (PETAL): a multicenter, randomized phase III trial. *J Clin Oncol*. 2018;36:2024-2034.