Patient age at diagnosis is associated with the molecular characteristics of diffuse large B-cell lymphoma

Wolfram Klapper¹, Markus Kreuz²*, Christian W Kohler³*, Birgit Burkhardt⁴*, Monika Szczepanowski¹*, Itziar Salaverria⁵*, Michael Hummel⁶, Markus Loeffler², Shoji Pellissery⁵, Wilhelm Woessmann⁴, Carsten Schwänen⁷, Lorenz Trümper⁸, Swen Wessendorf⁷, Rainer Spang³, Dirk Hasenclever², Reiner Siebert⁵, for the "Molecular Mechanisms in Malignant Lymphomas" Network Project of the Deutsche Krebshilfe

¹ Department of Pathology, Hematopathology Section and Lymph Node Registry, University Hospital Schleswig-Holstein, Campus Kiel/ University Kiel.

² Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Germany

³ Institute of Functional Genomics, University of Regensburg, Germany

⁴ NHL-BFM Study Center, Department of Pediatric Hematology and Oncology, Justus-Liebig-University, Giessen, Germany

⁵ Department of Human Genetics, University Hospital Schleswig-Holstein, Campus Kiel/ University Kiel, Germany

⁶ Institute of Pathology, Campus Benjamin Franklin, Charité–Universitätsmedizin Berlin, Germany
7 Cytogenetic and Molecular Diagnostics, Internal Medicine III, University Hospital of Ulm, Germany

8 Department of Hematology and Oncology, Georg-August University of Göttingen, Germany

* These authors made equal contributions

**Running title:** Patient age and diffuse large B-cell lymphoma

**Correspondence to:**

Prof. Wolfram Klapper, M.D., Department of Pathology, Hematopathology Section and Lymph Node Registry, University Hospital Schleswig-Holstein, Campus Kiel/Christian-Albrecht University Kiel. Michaelisstrasse 11, D-24105 Kiel, Germany

Phone: ++49 431 597-3399 or -4129

e-mail: wklapper@Path.uni-kiel.de
Abstract

Diffuse large B-cell lymphoma is the most frequent type of B-cell lymphoma in adult patients but also occurs in children. Patients are currently assigned to therapy regiments based on arbitrarily chosen age limits only (e.g. 18 or 60 years) and not biologically justified limits. A total of 364 diffuse large B-cell lymphomas and related mature aggressive B-cell lymphomas other than Burkitt lymphoma from all age groups were analyzed by comprehensive molecular profiling. The probability of several biologic features previously reported to be associated with poor prognosis in diffuse large B-cell lymphoma, such as ABC subtype, BCL2 expression, or cytogenetic complexity, increase with age at diagnosis. Similarly, various genetic features like IRF4 translocations, gains in 1q21, 18q21, 7p22 and 7q21 as well as changes in 3q27, including gains and translocations affecting the BCL6 locus, are significantly associated with patient age but no cut-offs between age groups could be defined. If age was incorporated in multivariate analyses, genetic complexity lost its prognostic significance, whereas the prognostic impact of ABC subtype and age were additive. Our data indicate that aging is a major determinant of lymphoma biology. They challenge current concepts regarding both prognostic biomarkers and treatment stratification based on strict age cut-offs.
Introduction

Diffuse large B-cell lymphoma (DLBCL) is a heterogeneous disease composed of different histopathological and genetic subtypes. Moreover, gene expression profiling has identified several molecular subgroups, of which the activated (ABC) and germinal center (GCB) subtypes of DLBCL are clinically the most relevant. Several studies have shown that the GCB subtype is associated with a more favorable prognosis than the ABC subtype. Additional molecular biomarkers for prognosis have been identified, including BCL2 protein expression, chromosomal aberrations and genetic complexity. However, these biomarkers for DLBCL were established almost exclusively for lymphomas in adults. Studies investigating biomarkers usually incorporated age as part of the clinical international prognostic index (IPI), because age is the strongest predictor of outcome in DLBCL. In the IPI age is used as a dichotomous variable with a cut-off at 60 years at diagnosis and higher age being an unfavorable risk factor.

DLBCL accounts for about 30% of adult lymphomas and 10% of lymphomas diagnosed before the age of 18 years. In clinical practice, children and adolescents up to the age of 18 years are usually treated according to protocols of pediatric lymphoma study groups, whereas patients older than 18 years are treated according to protocols of adult lymphoma study groups. Over the last decades, chemotherapy strategies in pediatric and adult DLBCL study groups have developed very differently, making a direct comparison of clinical results difficult. Nevertheless, the clinical outcome of DLBCL in children is much better than in adults. The treatment outcome has been greatly improved in adult patients by adding rituximab to CHOP-like regimens (cyclophosphamide, anthracyclins like doxorubicin, vincristine
and prednisolone every 21 or 14 days). The resulting survival rates for low-risk adult patients (who are predominantly young) are comparable to those for pediatric patients. The latter, however, are treated completely differently, with pediatric regimens which are short, 5 to 6 day dose-intensive courses including steroids, vincristine, high-dose methotrexate, cyclophosphamide or ifosfamide, doxorubicin, cytarabine, etoposide and intrathecal therapy for some risk groups.

Though the favorable prognosis of childhood and young adult DLBCL may be in part due to the differences in the treatment protocols or so far unknown host factors, several lines of evidence suggest that DLBCL in children differ biologically from their adult counterpart. First, pediatric DLBCL are almost exclusively of the centroblastic subtype; second, in childhood primarily the GCB subtype of DLBCL is observed; third, the chromosomal translocation t(14;18) involving the BCL2 gene, which is present in 20-30% of adult DLBCL of GCB subtype, is virtually absent before the age of 18 years, and finally, BCL2 or IRF4/MUM1 protein expression is found less frequently in children than adults. Overall, DLBCL in children seems to be molecularly more homogeneous than in adults, supporting the hypothesis of an age-dependent pathogenesis and biology. However, the age cut-off of 18 years used in clinical practice seems rather arbitrary and does not reflect DLBCL biology.

In line with the incidence pattern of DLBCL all large profiling studies published so far on this lymphoma have focused on adult and elderly patients and contain hardly any young adult or even pediatric patients. Thus, they do not allow us to conclude reliably whether differences in treatment strategies, in host characteristics or in tumor biology are key to the strong association between age and prognosis. Therefore, we
performed a comprehensive molecular characterization of DLBCL covering all age groups to determine possible age cut-offs, e.g. between pediatric and adult DLBCL, which might serve as a biological basis for assigning patients to treatment regiments.
Methods

Study Population

From a cohort of 742 samples from the network project "Molecular Mechanisms in Malignant Lymphoma" (MMML), 364 cases with DLBCL and related mature aggressive B-cell lymphomas other than Burkitt lymphoma were entered into this study based on a selection algorithm outlined in Figure S1. All biopsy specimens were evaluated by a panel of hematopathologists according to the WHO lymphoma classification \(^1\) by a process of independent analysis on single microscopes followed by discussion at a multi-head microscope to find a consensus for discrepant cases, as recently described.\(^5\) All cases were part of previous publications with a different focus.\(^5;21;22\). The study cohort included the histopathological diagnosis DLBCL, DLBCL in combination with a follicular lymphoma, atypical Burkitt lymphoma and aggressive mature B-cell lymphoma not otherwise specified. Molecularly defined Burkitt lymphomas (mBL) were excluded (see supplementary figure S1 for case selection). All biopsy specimens from obtained at relapse were excluded. For a detailed description of the study population see supplementary table 1 and supplementary figure 1. Due to the retrospective nature of the study and the broad range of age groups incorporated, the therapy applied was heterogeneous.\(^5;19\) The protocols of the MMML network were approved by ethics committees of all participating institutions. A previously published series of DLBCL homogeneously treated with R-CHOP was used as a control cohort (see supplementary figure 2B).\(^4\)

Immunohistochemistry

Immunohistochemical staining was performed on paraffin slides using standard techniques and antibodies against CD20, CD10, BCL2, BCL6, MUM1 and Ki67, as
described previously. The stainings were analyzed and scored semi-quantitatively by at least two observers. In cases where no paraffin-embedded material was available, the stainings were performed on frozen sections. For all analyses recently published guidelines for internal positive controls were applied.

Gene expression profiling and matrix (array) CGH

RNA and DNA were extracted from frozen sections (Qiagen, Hilden, Germany). Affymetrix U133A GeneChip (Santa Clara, CA, USA) hybridization was performed in accordance with the manufacturer’s recommendations using 5 μg of total RNA, as previously described. The gene expression data from the previously published series are available at http://www.ncbi.nlm.nih.gov/geo/ (Geo Accessions GSE4475, GSE10172 and GSE22470, respectively). The cell-of-origin signature (ABC, GCB), molecular Burkitt signature and PAP were assigned to each case based on the gene expression data.

Array-based comparative genomic hybridization was performed in 273 cases by applying a BAC/PAC array containing 2799 DNA fragments. The selection of recurrent copy number aberrations is described in the supplementary methods and was performed according to Kreuz et al.

Interphase fluorescence in situ hybridization

Interphase fluorescence in situ hybridization (FISH) was performed on frozen or paraffin-embedded tissue sections with the use of probes for IGH, IGK, IGL, MYC, BCL6, IRF4 and BCL2 loci. Tumor-biopsy specimens in which MYC was fused to IGH, IGK, or IGL (‘IG-MYC”) were distinguished from lymphomas with MYC breakpoints without fusion of MYC to an immunoglobulin locus (“non–IG-MYC”), as
previously described, but all lymphomas with MYC breaks were combined for the analysis in the current study.

EZH2 mutational analysis
Replacement of a single tyrosine in the SET domain of the EZH2 protein (Tyr641) was tested as previously described. In brief, Exon 15 of EZH2 was amplified by polymerase chain reaction using the primers EZH2_6812 (5'-tttgtcccatctccttc-3') and EZH2_6813 (5'-tggtcatttcccatca-3') and amplicons were subjected to direct sequencing using the same primers.

Statistical analyses
Differences in the incidence of biological features were analyzed by 2x2 tables and tested using Fisher’s exact test. This analysis was done to compare children (<15 years) and adults (>18 years). The conditional probability that a biological marker will be displayed, given the age at diagnosis, was analyzed using logistic regression.

In order to analyze the association between overall genomic instability and age a score for genetic complexity was designed. Genetic complexity was calculated as the sum of all detected genomic aberrations per lymphoma sample taking into account: t(14;18) IGH/BCL2 juxtaposition, BCL6 translocation, MYC translocation and all copy number aberrations listed in supplementary table 2. Correlation of age at diagnosis and genetic complexity was analyzed using Poisson regression. Overall survival was defined as time from first day of therapy to death from any cause. Patients without an event in OS were censored at the last day with valid information. Overall survival was estimated by the Kaplan-Meier method and compared using the log-rank test.
Multivariate analyses were done using Cox proportional-hazard models (see also supplementary material).
Results

Different molecular characteristics of pediatric and adult mature aggressive B-cell lymphoma other than Burkitt lymphoma

In order to determine which molecular features of DLBCL are associated with age a cohort of 364 mature aggressive B-cell lymphomas diagnosed as DLBLCL, composite FL/DLBCL or other high-grade B-cell lymphoma other than Burkitt lymphoma was analyzed for molecular features. In a first step, we stratified our cohort into children <15 y (n=15), adolescents of 15-18 y (n=5) and adults >18 y (n=344), as is commonly done in clinical practice, and tested for differences between children and adults in the incidence of biologic features of DLBCL. As already described, chromosomal translocations affecting the \textit{IRF4} locus were significantly more frequent in children than adults (p<0.001)\textsuperscript{22}. In contrast, gains of 2p16 and 6p25 were never detected in children, but in 17% and 15% of adult DLBCL, respectively (p=0.047 and p=0.048, supplementary Table 2 for all variables tested)

Continuous change of molecular characteristics in mature aggressive B-cell lymphoma (other than Burkitt lymphoma) with patient age In order to identify whether a clear cut-off can be defined between lymphomas derived from young (pediatric) and old (adult) patients based on the presence of molecular features can be defined, we plotted the different molecular variables according to the age at diagnosis (figure 1). Visual inspection of these plots suggested that several features additional to that identified in the comparison described above may have age-dependent frequencies of occurrence. Moreover, the increase in these features seems to be continuous throughout all age groups and appears to be strongest between 24-36 years rather
than 18 years (figure 1). However, the visualization has its pitfalls, because the incidence of lymphomas in elderly patients is much higher and the predominance of elderly patients in our cohort, which reflects the age distribution known for Germany (http://www.krebsregister-sh.de/), distorts the plot. Therefore, we modeled the conditional probability that a molecular feature will be displayed at a given age at diagnosis using logistic regression analysis (LRA), which eliminates the distorting effect of the underlying age distribution. This analysis determines the likelihood that a lymphoma will show a molecular feature in relation to the patient age at the time of diagnosis. LRA revealed a statistically significant association between increasing age and the cell-of-origin signature of the ABC subtype, BCL2 protein expression, absence of IRF4 translocations, gains in 1q21, 18q21, 7p22 and 7q21 as well as changes in 3q27, including gains and translocations affecting the BCL6 locus (figure 2A and supplementary table 3 for all variables tested). In fact, the genetic complexity, measured as the sum of detectable genomic aberrations, increased continuously with age (figure 2B). LRA confirmed that there is a significant continuous increase in the ABC probability not only in our data set but also in an independent case collection (p=0.042) (supplementary figure 2). Poisson regression showed that the number of chromosomal aberrations in a lymphoma increased significantly with age (Fig. 2A) with the shape of the curves of the LRA analyses and the Poisson regression suggesting a continuous increase in the probability of molecular features with patient age rather than a strict border between age groups (Fig. 2).

Interdependence of age and molecular characteristics in relation to prognosis.
We tested for prognostic significance in a Cox model combining each molecular feature that was identified by LRA, with patient age. As table 1 shows, the prognostic
impact of the molecular ABC subtype (p=0.0001) and BCL2 protein expression (p=0.0019) was maintained in a Cox model including age. In contrast, all genetic markers lost the significant prognostic impact if age was taken into account. We further validated the prognostic significance of the molecular subtype in a second and completely independent cohort homogeneously treated with immunochemotherapy and confirmed the prognostic information of age and ABC/GCB subtype to be independent and additive (supplementary Fig. 2B). Further molecular data, such as BCL2 protein expression, were not accessible for analysis in this cohort.
Discussion

The age cut-off of 18 years is commonly used to stratify patients for pediatric and adult therapy regimens. However, therapeutic approaches to DLBCL in pediatric and adult lymphoma study groups have developed very differently over the last decades. Children and adolescents up to the age of 18 years are treated with regimens that are also applied for Burkitt lymphoma, whereas anthracyclin based therapy in combination with anti-CD20 antibody is the current gold standard for adult patients.12 This study showed that molecular features of DLBCL are associated with age and that prognostically unfavorable molecular features like the ABC subtype and BCL2 expression increase with patient age at diagnosis. Our data are in line with previously published immunophenotypes of pediatric DLBCL.30 However, our data do not support a pathogenetic dichotomy between DLBCL in children and adults (“two lymphoma model”), as initially hypothesized. For all features associated with increasing age (ABC subtype, BCL2 protein expression, absence of IRF4 translocations, gains in 1q21, 18q21, 7p22 and 7q21 as well as changes in 3q27, including gains and translocations affecting the BCL6 locus, genetic complexity), there was no clear age cut-off between pediatric and adult DLBCL. Remarkably, the genetic features that increased in likelihood with age in the LRA included changes associated with both ABC (changes 3q27/BCL6, +18q21) and GCB type DLBCL (+1q21, +7p22, +7q21),31 suggesting that there is an increase in genetic complexity with age that is to some extent independent of the molecular lymphoma subtype. The continuous increase in genomic complexity with age might suggest an “age evolution model” characterized by a stochastic risk of acquiring genetic aberrations with age. Probably the type of genetic changes that accumulated in a lymphoma initiating cell determines the manifestation of a certain molecular subtype of DLBCL. In relation to
the choice of therapy, our data strongly argue for the development of common treatment protocols for DLBCL in adolescents and young adults since there is no biological rationale for a clearly definable age limit and the therapy would be tolerated equally well.

Interestingly, differences in molecular features of pediatric and adult disease are not restricted to DLBCL, but have recently been demonstrated for lymphoblastic leukemia \(^{32}\) and follicular lymphoma.\(^ {33}\) However, despite differences in clinical presentation, such as gender distribution, we have failed to demonstrate molecular differences between pediatric and adult Burkitt lymphoma so far.\(^ {19}\) Nevertheless, molecular characterization of malignant tumors occurring in children and adults might be a very fruitful approach for gaining insights into disease biology.

The clinical course after chemotherapy for adult DLBCL varies substantially between individual patients. As a consequence, efforts have been made to establish prognostic biomarkers by analyzing lymphoma biology. However, the IPI, which is based on the clinical features age, lactate dehydrogenase value, performance status, Ann Arbor stage and number of extranodal involvements, still seems to be the most powerful prognostic tool in DLBCL.\(^ {34}\) In the IPI age is reflected by categorizing the patients as younger or older than 60 years of age. Most studies of biomarkers correlate with the IPI as the current prognostic gold standard. The cohort presented here is of limited value for survival analysis due to the heterogeneous treatment applied. Nevertheless, our data suggest that in future studies prognostic biomarkers analyzed in homogeneously treated cohorts should be very carefully analyzed for age association. Since the age at disease onset influences the biology of the
lymphoma, prognostic markers might lose their prognostic significance if age is depicted in survival analysis as a continuous parameter. However, we also demonstrated that the prognostic significance of the molecular subtypes of DLBCL is independent of the patient age.

In summary the present study, which represents the largest cohort of DLBCL analyzed by comprehensive molecular profiling so far, challenges current pathogenetic and clinical concepts in DLBCL. It questions the biologic rationale for the use and level of age cut-offs for stratification in clinical trials as well as the application of some age-associated biomarkers.\textsuperscript{32,35}

Acknowledgements

The study was supported by the Deutsche Krebshilfe Network (70-3173-Tr3) “Molecular Mechanisms in Malignant Lymphomas” and the ICGC MMML-Seq of the German Ministry of Education and Science (BMBF, No. 01KU1002A). WK and RS are supported by the Kinderkrebsinitiative Buchholz, Holm-Seppensen. IS is supported by a fellowship from the Alexander von Humboldt Foundation. The funding sources did not influence study design, data collection, analysis, interpretation, writing of the report, and in the decision to submit the paper for publication. The authors would like to thank Olivera Batic, Claudia Becher, Reina Zühlke-Jenisch and Charlotte Botz von Drathen for their excellent technical support.

Conflict of Interests

The authors have no competing financial interests to declare.
Author contributions

M.K., C.K., M.L., D.H., and R.Sp. performed the statistical analysis; M.S., I.S., M.H., C.S., S.P. and S.W. performed the molecular analysis and analyzed data; B.B., W.W. and L.R. provided samples and clinical data; W.K. and R.S. designed the project, analyzed the data, and wrote the manuscript.
Reference List


Table 1: Prognostic impact of features associated with age in an univariate analysis and after adjustment for age. * Genetic complexity as a continuous variable. GE= gene expression, IHC= immunohistochemistry.
FIGURE LEGENDS

Figure 1
Scatter plot showing the phenotypic variables on the x-axis, with each dot being an event-positive representative at its corresponding age level (y-axis: age at diagnosis in years). The variables are not exclusive, since a case with an IgH-break might also be displayed in as being IgH-BCL2 translocation positive. Note that this visualization is distorted by the skewed distribution of age at diagnosis. The color code indicates membership of each variable to an individual group: green = gene expression, red = chromosomal translocation/breaks, blue = chromosomal copy number gains, orange = chromosomal copy number losses, magenta = immunohistochemistry markers, gray = morphology and black = EZH2 Tyr 641 mutation status. The span from 24-36 years in which visual inspection reveals a strong increase in several molecular features is highlighted in gray.

Figure 2
A) Logistic regression analyses of ABC/GCB gene expression groups, immunohistochemical BCL2 expression, IRF4 translocations, 1q21 gains, changes in 3q27 (including gains and translocations of BCL6), 18q21 gains, 7p22 gains and 7q21 gains. Patients with the respective feature are plotted in red scattered around the horizontal at 1, patients lacking the respective feature are given in green at the bottom scattered around 0. The dashed line indicates the overall frequency of the feature irrespective of age. The blue logistic regression curve shows the estimated conditional probability of having the feature at a given age at diagnosis.
B) Scatter plot of age at diagnosis (in years) versus genetic complexity calculated as the sum of all detected genomic aberrations per lymphoma sample taking into account: t(14;18) IGH/BCL2 fusion, BCL6 translocation, MYC translocation and all copy number aberrations listed in supplementary table 2. The line indicates the Poisson regression curve representing the estimated average number of abnormalities as a function of age at diagnosis.
Figure 2

A

B

Genetic complexity as a function of age

Number of genetic abnormalities (filtered) vs. age

P-value = 0.0000007
Patient age at diagnosis is associated with the molecular characteristics of diffuse large B-cell lymphoma

Wolfram Klapper, Markus Kreuz, Christian W Kohler, Birgit Burkhardt, Monika Szczepanowski, Itziar Salaverria, Michael Hummel, Markus Loeffler, Shoji Pellissery, Wilhelm Woessmann, Carsten Schwänen, Lorenz Trümper, Swen Wessendorf, Rainer Spang, Dirk Hasenclever and Reiner Siebert