5-2022

Diffuse Large B-Cell Lymphoma Risk Score Calculator: A Tool To Accurately Predict Survival Risk And Improve Treatment Decisions And Outcomes

Santosh Khanal
Children's Mercy Hospital

Arjun Gampala
Children's Mercy Kansas City

Eric S. Geanes
Children's Mercy Hospital

Anne T. Culpepper

Liubomyr Buryi

See next page for additional authors

Authors

Santosh Khanal, Arjun Gampala, Eric S. Geanes, Anne T. Culpepper, Liubomyr Buryi, Catherine A. Jackson, Rebecca McLennan, and Todd Bradley

DIFFUSE LARGE B-CELL LYMPHOMA RISK SCORE CALCULATOR: A TOOL TO ACCURATELY PREDICT SURVIVAL RISK AND IMPROVE TREATMENT DECISIONS AND OUTCOMES

Santosh Khanal[1], Arjun Gampala[1,2], Eric S. Geanes[1], Anne T. Culpepper[3], Liubomyr Buryi[3]
Catherine A. Jackson[4], Rebecca McLennan[1], Todd Bradley[1,5,6]

[1] Center for Pediatric Genomic Medicine, Children's Mercy Kansas City. [2] University of Missouri Kansas City Medical School. [3] illumisoft. [4] Research Informatics, Children's Mercy Kansas City. [5] Department of Pediatrics, University of Missouri Kansas City Medical School. [6] Departments of Pediatrics and Pathology and Laboratory Medicine, University of Kansas Medical Center.

Abstract

Diffuse large B-cell lymphoma (DLBCL) is the most common type of non-Hodgkin lymphoma. It is an heterogenous cancer that can have variable responses to therapy and profound differences in survival outcomes. DLBCL is of B- cell origin and was typically treated with a regimen of cyclophosphamide, hydroxydaunorubicin, oncovin and prednisone (CHOP) but the addition of the anti-CD20 monoclonal antibody rituximab (R) significantly improved patient overall-survival (OS). A scoring system has been used to identify risk groups of DLBCL individuals called the International Prognostic Index (IPI) that uses age, lactate dehydrogenase levels, general health status, stage of tumor and number of disease sites to place the patients in 1 of 4 risk groups that correspond with the likelihood of 3-year OS. A revised IPI (R-IPI) using R-CHOP-treated patients was developed that had improved prognostic value at determining risk groups. What this historically used DLBCL scoring system lacks is more individualized data, including molecular information. Limited gene expression studies have been performed on DLBCL, which have successfully classified DLBCL subclasses. Here, we have expanded on previous gene expression studies and have identified genes that are associated with OS in DLBCL using data from the Lymphoma/Leukemia Molecular Profiling Project. Specifically, we have developed a prognostic gene signature consisting of 33 genes that, when transformed into a risk score, can stratify individuals into high or low risk groups that have significantly different OS. The prognostic gene signature was associated with OS in multiple clinical studies, and when used in conjunction with DLBCL molecular subtype and R-IPI score, OS was predicted more accurately. Recently, we have developed a web based DLBCL Risk Score Calculator. This calculator combines our prognostic gene signature and the current R-IPI risk scoring system to generate risk scores and predict survival probabilities by year. Next, we will challenge our DLBCL Risk Score Calculator by performing RNA sequencing on DLBCL samples with documented pathology and progress reports. After validation, our DLBCL Risk Score Calculator will more accurately identify patients that will do well on standard care, and alert clinicians to patients that would benefit from experimental therapeutic approaches or more frequent clinical monitoring following treatment.

Background

Diffuse Large B Cell Lymphoma (DLBCL) is a neoplasm arising from the lymphoid progenitor cell line, and it is the most common lymphoma in the United States. There are 2 subclasses, including the germinal center B cell-like (GCB) and the activated B cell-like (ABC), with the ABC class being more aggressive. DLBCL frequently presents with lymphadenopathy and hepatosplenomegaly but can be diagnosed based on molecular markers including CD19, CD20, CD10, BCL6, and CD5. There is a widespread incidence of DLBCL in middle and older aged adults, and currently, the methods of staging and treatment strategies do not rely on new genetic based technologies. The Revised International Prognostic Index (R-IPI) is used by physicians to identify risk groups based on age, lactate dehydrogenase levels, general health status, stage of tumor and number of disease sites. However, by examining the expression levels of several genes in DLBCL patients, we may be able to create a genetic based risk score calculator that more accurately predicts overall survival (OS) in patients. This will allow physicians to describe disease progression as well as evaluate the benefit of treatment.

Purpose

To create a genetic based risk calculator for physicians that more accurately predicts the overall survival of DLBCL patients and enhances clinical decision making.

Experimental Approach

1. Identify genes associated with DLBCL

- We identified 233 DLBCL samples from the Genome Expression Omnibus (under the accession number GSE10846) that had established gene expression values
- We categorized the genes into high or low risk expression based on the median gene expression value
- Using the R survival package, we plotted the Kaplan Meier curves and found the p-values for each gene
- 61 genes were found to be statistically significant ($p<0.001$)
- We performed LASSO (Least Absolute Shrinkage and Selection Operator) analysis on the 61 statistically significant genes
- 33 genes were found to have non-zero coefficients
- By multiplying these coefficients with a patient's expression values of the 33 genes we identified and then adding those values together, we created a risk score
- We validated our risk score with 3 external datasets (GSE34171, GSE32918/69051, and TCGA)

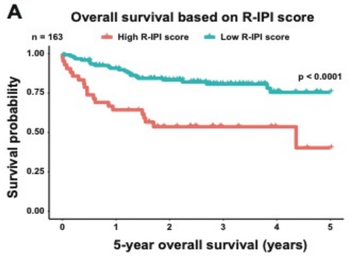
2. Develop DLBCL Risk Score Calculator

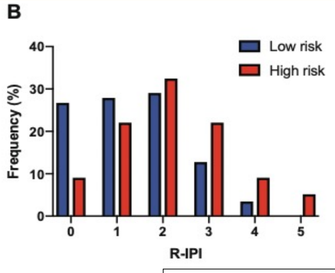
- Together with Illumisoft, we developed this DLBCL risk score calculator into an app and website
- The web app leverages a single page architecture with dynamic routing and is written in the React framework
- Hosted on Azure Cloud's app services

3. Sequence RNA from DLBCL samples

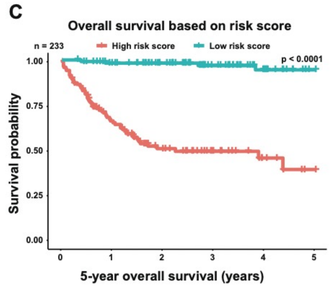
- 10 DLBCL patient FFPE samples were obtained from the Biospecimen Repository Core Facility The University of
- After deparaffinization, we performed Bulk RNA sequencing on the deparaffinized FFPE tissue
- We evaluated the gene expression values for the 33 genes we identified for our risk score calculator
- We calculated the risk scores for these patients and compared them to R-IPI

The Prognostic Gene Signature More Accurately Places Patients into Risk Groups



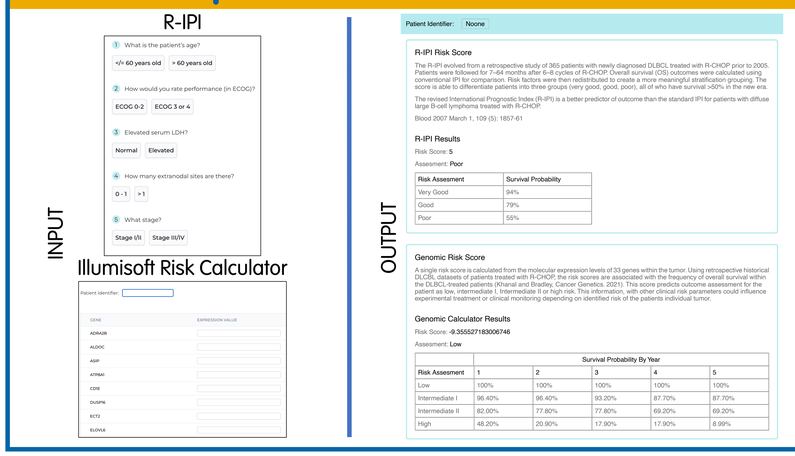


When comparing our prognostic gene signature to the R-IPI, we found that those who had high risk scores had significantly lower overall survival than previously assigned by R-IPI, and those who had a low-risk value had higher overall survival. As shown in these graphs, there was a greater margin between the high risk and low risk assigned mortalities in the prognostic gene signature.

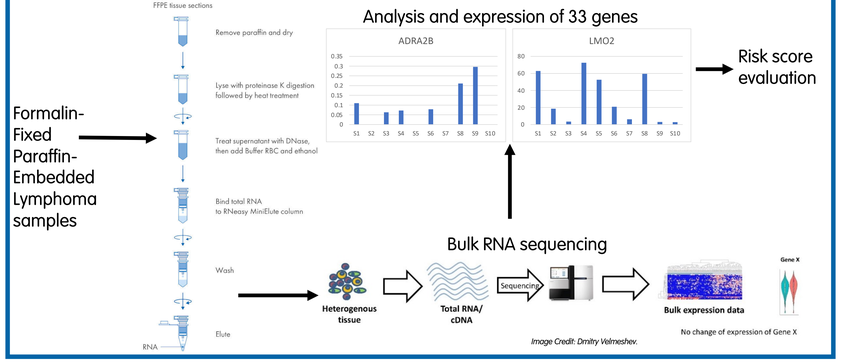


Gene name	Lasso coefficient
ADRA2B	0.05929974
ALDOC	-0.2266974
ASIP	-0.0994086
ATP8A1	-0.052468
CD1E	-0.1111254
DUSP16	-0.0963421
ECT2	0.13182723
ELOVL6	0.055146
FAF1	-0.0652772
FAM223A|FAM223B	-0.0121265
CAREM	-0.0295263
GNG8	-0.0089058
IGSF9	0.19446142
LMO2	-0.0070721
LPP84	-0.1433395
LY75	-0.252489
MAEL	-0.086909
NEK3	0.08073014
PADI2	-0.0332634
PDK1	-0.0435511
PDK4	0.18311325
PES1	0.09271489
PPP1R7	-0.2483229
PUSL1	0.14247471
SCN1A	-0.054923
SLAMF1	-0.0094785
SSTR2	-0.0260066
TADA2A	0.12055065
TNFRSF9	-0.004922
USH2A	-0.1920536
VEZF1	-0.3893348
WDR91	-0.0041198
ZMYND19	0.26520514

Development of Risk Score Calculator



Current DLBCL Patient Sample RNA Analysis



Summary

- We developed a 33 prognostic gene signature that accurately identifies overall survival and can be used in conjunction with the current regimen of DLBCL care.
- Our prognostic gene signature predicted better OS for 3 external datasets (GSE34171, GSE32918/69051, and TCGA) than what R-IPI initially determined.
- We are currently validating our prognostic gene signature by testing it on DLBCL patient samples and determining their risk scores
- For future studies, a larger sample size of DLBCL patients to validate the risk calculator may strengthen the accuracy
- Also, using a larger sample size from databases may eliminate the differences from evaluating RNA data attained by different techniques in the various databases.
- Together, R-IPI and our prognostic gene signature can help physicians determine the prognosis of their patients as well as make therapeutic decisions for their patients

Acknowledgements

Lymphoma FFPE samples were supplied by the Biospecimen Repository Core Facility, The University of Kansas Cancer Center. Funding for this work was through internal institutional funds to T.B. from Children's Mercy Research Institute and Children's Mercy Kansas City, in particular an innovator award from Mercy Research Partners.

UMKC School of Medicine
KU School of Medicine The University of Kansas

Laboratory of Immunogenomics
at Children's Mercy Kansas City

Children's Mercy Kansas City | Research Institute