ELSEVIER

Blood 142 (2023) 4353–4354

The 65th ASH Annual Meeting Abstracts

POSTER ABSTRACTS

618.ACUTE LYMPHOBLASTIC LEUKEMIAS: BIOMARKERS, MOLECULAR MARKERS AND MINIMAL RESIDUAL DISEASE IN DIAGNOSIS AND PROGNOSIS

Epigenetic Profiling and Machine Learning for Enhanced Risk Stratification in Pediatric Acute Lymphoblastic Leukemia

Adrian Mosquera Orgueira[1], *Olga Krali*[2], *Andres Peleteiro Raindo*[3], *Jose Angel Diaz Arias*[4], *Marta Sonia Gonzalez Perez*[5], *Manuel Perez Encinas*[6], *Manuel Fernandez Sanmartin*[7], *Natalia Alonso Vence*[8], *Daniel Sinnett, PhD*[9], *Mats Heyman*[10], *Gudmar Lönnerholm*[11], *Ulrika Norén-Nyström, MD PhD*[12], *Kjeld Schmiegelow*[13], *Jessica Nordlund*[2]

[1] University Hospital of Santiago de Compostela, Department of Hematology, IDIS, SANTIAGO DE COMPOSTELA, Spain
[2] Department of Medical Sciences and Science for Life Laboratory, Uppsala University, Uppsala, Sweden
[3] University Hospital of Santiago De Compostela, Department of Hematology, IDIS, SANTIAGO DE COMPOSTELA, ESP
[4] University Hospital of Santiago de Compostela, Department of Hematology, IDIS, A Coruña, Spain
[5] University Hospital of Santiago de Compostela, Santiago de Compostela, Spain
[6] University Hospital of Santiago de Compostela, Department of Hematology, IDIS, SANTIAGO, ESP
[7] University Hospital of Santiago de Compostela, Department of Pediatrics, IDIS, Santiago de Compostela, Spain
[8] University Hospital of Santiago de Compostela, Department of Hematology, IDIS, Santiago de Compostela, Spain
[9] Hopital Ste.-Justine Centre de Cancerologie Charles Brunea, Montreal, CAN
[10] Childhood Cancer Research Unit, Karolinska Institute, Stockholm, Sweden
[11] Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden
[12] Department of Clinical Sciences, Pediatrics, Umeå University Hospital, Umeå, Sweden
[13] Copenhagen University, Copenhagen, DNK

Introduction:
Acute lymphoblastic leukemia (ALL) is the most common childhood malignancy with high clinical and biological heterogeneity. Despite advancements in treatment, relapsed patients still face significant mortality rates and long-term complications. Risk-adapted strategies have been implemented to optimize treatment intensity based on prognostic factors. Genomic techniques offer potential for improved risk stratification, but accurate grouping remains challenging. Epigenetic alterations, particularly DNA methylation, have shown promise in tumor classification and prognostication. In this study, we trained machine learning survival models based on DNA methylation signatures to refine risk grouping in pediatric ALL patients.

Methods:
Clinical annotation and DNA methylation data from pediatric ALL samples were retrieved from the *Norlund et al.* cohort (n = 763). Age, sex, risk group, and cytogenetic subtype were selected as clinical covariates. The cohort was randomly divided into training (80%) and test (20%) sets. Univariate Cox regression and variable importance analysis were performed to select CpG sites associated with relapse-free survival (RFS) and overall survival (OS). Random survival forest models were constructed using the training set, and their performance was evaluated on the test set using the concordance index (c-index), the time-dependent area under the ROC curve (AUC), and the continuous rank probability scores (CRPS). To assess the generalizability of the risk predictors, we performed external validation using two independent pediatric ALL datasets (*Busche et al.*, n = 42; and *Krali et al.*, n = 384).

Results:
The relapse risk predictor (RRP) was constructed using random survival forests based on a signature of 16 CpG sites. The model achieved good predictive accuracy with c-indexes of 0.667 and 0.677 in the training and test sets, respectively. The addition of cytogenetic subtype or age at diagnosis did not significantly change the model's performance. Longitudinal assessment of the RRP revealed its superior performance compared to clinical risk grouping. Combining the RRP with clinical risk grouping improved prognostic accuracy, with a 20-month AUC over 80%.
The mortality risk predictor (MRP) was constructed using a signature of 53 CpG sites. The model achieved strong predictive performance with c-indexes of 0.751 and 0.755 in the training and test sets, respectively. Similar to the RRP, the addition of cytogenetic subtype or age at diagnosis did not significantly impact the model's performance. Longitudinal assessment of

the MRP demonstrated a higher performance rate than clinical risk grouping at all evaluated time points. Combining the MRP with clinical risk grouping yielded the highest prognostic accuracy.
In the external validation of the MRP in the *Krali* dataset, the MRP score was strongly associated with OS: c-index 0.621, p-value 1.06×10^{-4}. The hazard ratio was 1.073 (95% confidence interval: 1.035-1.112) for each incremental increase in the risk score. On the contrary, a lower reproducibility of the RRP was observed (c-index 0.529), presumptively related to the addition of MRD-driven risk stratification in more recent protocols.

Conclusions:
Machine learning models built on DNA methylation signatures surpassed the traditional clinical risk grouping both for predicting relapse risk and mortality, while the combination of molecular and clinical factors provided the best prognostic accuracy. Further validation and implementation of these predictors could contribute to personalized risk-adapted treatment strategies for pediatric ALL patients.

Disclosures Mosquera Orgueira: *Janssen:* Consultancy; *AstraZeneca:* Consultancy.

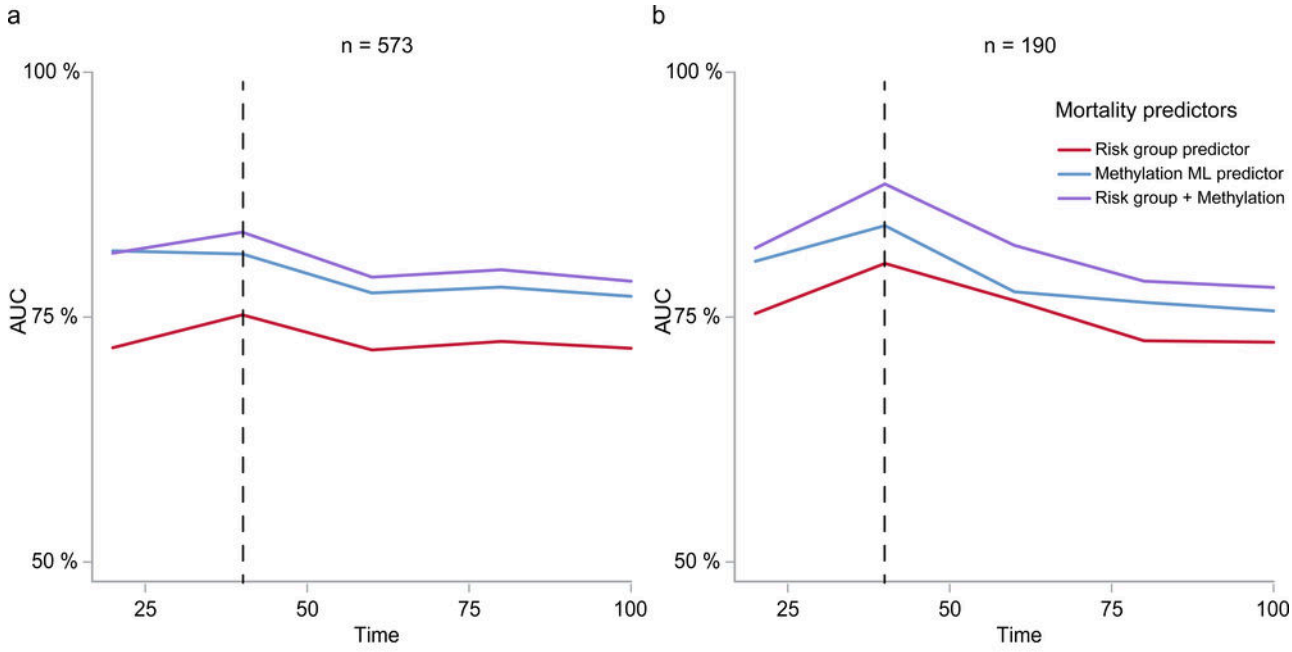

Figure 1. Time-dependent area under the curve (AUCs) as performance metrics for three mortality risk classifiers for the training (a) and test (b) sets. The model based on standard risk groups is denoted with red color, the mortality risk predictor (MRP) with blue and the models integrating both risk grouping and MRP with purple.

Figure 1

https://doi.org/10.1182/blood-2023-179156