5021

cfDNA Methylation-Based Pan-Hematologic Prognostic Classification

ASH 2023
December 9-12, 2023
San Diego, CA, USA

Yuefan Huang, PhD, Alvin Shi, PhD, Rita Shaknovich, MD, PhD, Oliver Venn, PhD, Qinwen Liu, PhD
GRAIL, LLC, Menlo Park, CA, USA

INTRODUCTION

- Accurate prognosis prediction is important to guide effective treatment strategies in hematologic (heme) malignancies
- However, determining the prognosis of heme malignancies remains challenging because of the complex and diverse biology of these diseases and the invasive nature of prognostic procedures (eg, bone marrow aspiration and biopsy)[1]
- Risk stratification and prognostic methods (eg, cytogenetics, gene mutation panels, clinical parameters)[2] exist for certain single heme indications, but in light of changing classifications of heme neoplasms, many indications lack prognostic tools[3-7]
- A minimally invasive pan-heme (encompassing all hematologic lineages) prognostic tool could efficiently fill this clinical need
- Targeted methylation (TM) sequencing of cell-free DNA (cfDNA) provides a means of capturing the biological heterogeneity of multiple heme neoplasms using a simple, single blood draw as a liquid biopsy
- We leveraged GRAIL's multi-cancer early detection platform that utilizes TM sequencing of cfDNA from peripheral blood and large population-level datasets to train a machine learning-based pan-heme classification model to predict prognosis and better inform clinical management of patients

OBJECTIVE

- Demonstrate proof-of-concept performance of a pan-heme, methylation-based prognostic classification model, focusing on one representative indication per major heme lineage

METHODS

Participant Selection

- We selected participants from the Circulating Cell-free Genome Atlas substudy 2 (NCT02889978)[8] who were diagnosed with specific hematologic malignancies/neoplasms (**Table 1**). Participants with prior cancer history or with a cancer subtype that had < 5 total observations across the study were excluded
- 636 participants were selected for model training, and 466 participants were used for model evaluation with 6-fold cross-validation (**Table 1**)

Sample Processing

- cfDNA from plasma samples was extracted and subjected to bisulfite conversion, enrichment, and TM sequencing through a previously described proprietary GRAIL protocol (GRAIL, LLC)[8]

Methylation Prognostic Model Training and Evaluation

- To train prognostic models, we used a Random Survival Forests (RSF)[9] model to predict overall survival (starting from the time of blood draw)
- The baseline clinical characteristics-only prognostic model integrated clinical variables, including heme cancer subtype, race, age, sex, stage, BMI, smoking status, and drinking status
- The methylation-based prognostic model included additional features based on methylation principal components (PC) computed from the beta-value matrix
- For both models, hyperparameter optimization (number of PCs, mtry [the number of features to consider at each split point], node size) was conducted under a 6-fold nested cross-validation framework
- The concordance index (C-index) was used to assess the models' predictive performance, and confidence intervals were generated using 100 bootstrap samples

Heme Subtype Risk Stratification

- The risk cutoff for each heme neoplasm subtype was determined by the corresponding 6 outer training fold scores. The optimal risk score thresholds were determined with Youden's index (Youden's J statistic; J = sensitivity + specificity − 1)
- Cutoff values were derived by applying Youden's index to 5-year survival Receiver Operator Characteristic (ROC)
- Using Youden's index, participants were stratified into 2 distinct groups (high- and low-risk) based on the predicted risk scores

Survival Analysis

- To investigate if prognostic classifiers were able to stratify participant risk, we used the log-rank test to examine survival differences and Kaplan-Meier plots to visually represent the survival distribution across different risk groups

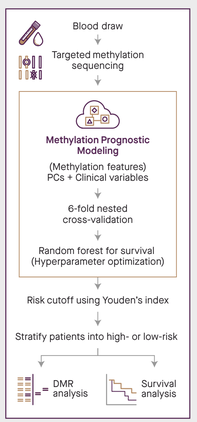
Differential Methylation Analysis

- Differential methylation analysis of cfDNA samples from high-risk and low-risk participant groups was performed using beta-binomial regression with an arcsine link function to model the region-level count data, adjusting for sex.[10] Hypothesis testing was performed using a Wald's test for each region
- The differentially methylated regions (DMRs) were considered significant if the q-value was < 0.05 and the absolute value of beta-binomial model delta coefficients was > 0.1
- Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis[11] was performed using the closest genes to each significant hyper- or hypo-DMR using a hypergeometric test

Table 1. Participant Characteristics.

	Training samples N=1163	Training participants N=636	Evaluation samples N=913	Evaluation participants N=466
Heme neoplasms (% of total)				
Circulating lymphomas (eg, CLL, SLL)	118 (10%)	65 (10%)	99 (11%)	50 (11%)
Hodgkin lymphoma	56 (4.8%)	32 (5.0%)	34 (3.7%)	18 (3.9%)
MGUS	241 (21%)	124 (19%)	205 (22%)	105 (23%)
Myeloid neoplasms (eg, polycythemia vera)	275 (24%)	145 (23%)	218 (24%)	112 (24%)
Non-Hodgkin lymphomas (eg, B cell lymphoma, DLBCL, follicular lymphoma, MALT/NMZL, mantle cell)	357 (31%)	198 (31%)	271 (30%)	137 (29%)
Plasma cell neoplasm	116 (10%)	72 (11%)	86 (9.4%)	44 (9.4%)
Sex (% of total)				
Female	560 (48%)	304 (48%)	433 (47%)	220 (47%)
Male	603 (52%)	332 (52%)	480 (53%)	246 (53%)
Age, mean (SD)	64.37 (13.61)	64.32 (13.57)	64.39 (13.35)	64.38 (13.38)
Race (% of total)				
American Indian or Alaska Native	7 (0.6%)	4 (0.6%)	6 (0.7%)	3 (0.6%)
Asian, Native Hawaiian or Pacific Islander	15 (1.3%)	8 (1.3%)	14 (1.5%)	7 (1.5%)
Black, non-Hispanic	85 (7.3%)	46 (7.2%)	71 (7.8%)	36 (7.7%)
Hispanic	113 (9.7%)	60 (9.4%)	84 (9.2%)	43 (9.2%)
Other/missing	6 (0.5%)	3 (0.5%)	6 (0.7%)	3 (0.6%)
Other/unknown	7 (0.6%)	5 (0.8%)	6 (0.7%)	4 (0.9%)
White, non-Hispanic	930 (80%)	510 (80%)	726 (80%)	370 (79%)
Overall survival, mean days (SD)	1272.15 (439.14)	1284.91 (446.14)	1213.94 (416.51)	1211.96 (416.27)
Overall survival, events (% of total)	189 (16%)	108 (17%)	138 (15%)	71 (15%)

CLL, chronic lymphocytic leukemia; DLBCL, diffuse large B cell lymphoma; heme, hematologic; MALT, mucosa-associated lymphoid tissue lymphoma; MGUS, monoclonal gammopathy of unknown significance; NMZL, nodal marginal zone lymphoma; SD, standard deviation; SLL, small lymphocytic lymphoma.

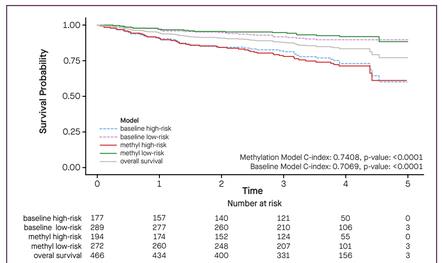


KEY RESULTS: A NOVEL BLOOD-BASED, PAN-HEME PROGNOSTIC CLASSIFICATION MODEL DEMONSTRATED POTENTIAL UTILITY FOR RISK STRATIFICATION ACROSS MULTIPLE HEME NEOPLASMS

Methylation Data Offer Superior Prognostic Performance Over a Baseline Clinical Characteristics-Only Model

- We assessed the performance of a methylation-based prognostic classification model ("methylation prognostic model") and compared it to a baseline clinical characteristic-only prognostic model to assess the additional effect of methylation data on prognostic performance
- The methylation prognostic model demonstrated better discrimination than the baseline clinical characteristics-only prognostic model for pan-heme neoplasms in terms of C-index, a measure of predictive performance (**Figure 1**)
  - Methylation prognostic model overall C-index: 0.7408 (95% confidence interval [CI]: 0.6759-0.7960)
  - Baseline clinical characteristics-only prognostic model overall C-index: 0.7069 (95% CI: 0.6438-0.7650)
- When heme neoplasm subtypes were assessed individually, the methylation prognostic model demonstrated better discrimination than the baseline clinical characteristics-only prognostic model for 9 subtypes (**Figure 2**)
- For follicular lymphoma, polycythemia vera, and plasma cell neoplasm, the methylation prognostic model demonstrated statistically significant risk stratification when the baseline clinical characteristic-only prognostic model did not (**Figure 3**)
  - Other subtypes also showed promising stratification; however, the interpretations are preliminary due to limited sample sizes

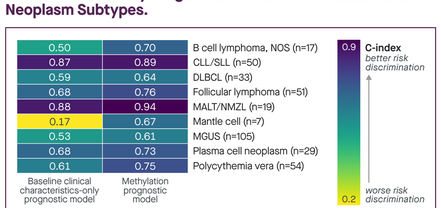
Figure 1. The Methylation Prognostic Model Demonstrated Better Pan-Heme Risk Discrimination Than the Baseline Clinical Characteristics-Only Prognostic Model in Term of C-Index.



Kaplan-Meier survival curves for pan-heme risk stratification by the methylation prognostic model (solid lines) and baseline clinical characteristics-only prognostic model (dashed lines). C-index was used to evaluate the predictive performance of the models to discriminate risk (a C-index of 1 represents perfect discrimination). p-values indicate the statistical significance of the log-rank test between the high- and low-risk groups. The overall survival of the population cohort, unadjusted for any risk factors, is represented by a gray line.

C-index, concordance index; heme, hematologic.

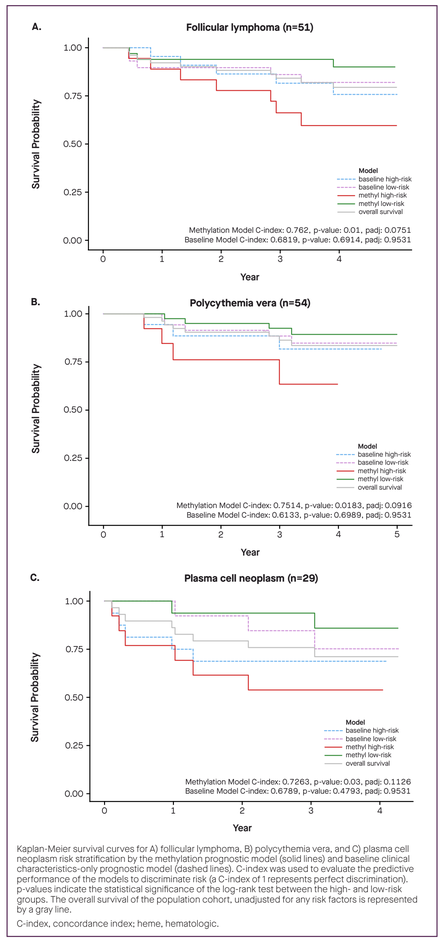
Figure 2. The Methylation Prognostic Model Demonstrated Equal or Better Risk Discrimination Than the Baseline Clinical Characteristics-Only Prognostic Model for Individual Heme Neoplasm Subtypes.



Risk prediction performance (evaluated by C-index) of the methylation prognostic model and the baseline clinical characteristics-only prognostic model for 9 heme neoplasm subtypes is visualized as a heatmap. C-index measures the models' ability to discriminate risk, with higher C-index values indicating better discrimination. The methylation prognostic model had the best performance in these 9 subtypes out of 14 total subtypes evaluated. The number of evaluable participants per subtype is included in parentheses.

C-index, concordance index; CLL/SLL, chronic lymphocytic leukemia/small lymphocytic lymphoma; DLBCL, diffuse large B cell lymphoma; heme, hematologic; MALT, mucosa-associated lymphoid tissue lymphoma; MGUS, monoclonal gammopathy of unknown significance; NMZL, nodal marginal zone lymphoma; NOS, not otherwise specified.

Figure 3. The Methylation Prognostic Model Demonstrated Superior Risk Stratification Relative to the Baseline Clinical Characteristics-Only Prognostic Model for Individual Heme Neoplasm Subtypes.



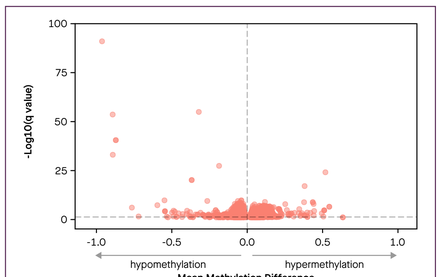
Kaplan-Meier survival curves for A) follicular lymphoma, B) polycythemia vera, and C) plasma cell neoplasm risk stratification by the methylation prognostic model (solid lines) and baseline clinical characteristics-only prognostic model (dashed lines). C-index was used to evaluate the predictive performance of the models to discriminate risk (a C-index of 1 represents perfect discrimination). p-values indicate the statistical significance of the log-rank test between the high- and low-risk groups. The overall survival of the population cohort, unadjusted for any risk factors is represented by a gray line.

C-index, concordance index; heme, hematologic.

DMR Analysis Revealed Key Biologic Pathways Distinguish High- and Low-Risk Participants

- DMRs were defined as regions within cfDNA that had significantly different methylation states between training set high- and low-risk participant groups. Across all heme neoplasm subtypes, 12,005 significant DMRs were identified between the high- and low-risk groups (**Figure 4**)
- DMR analysis highlighted extensive epigenetic alterations, predominantly in cancer-related signaling pathways, such as Wnt, MAPK, and Hippo. These pathways were identified in both hyper-methylated and hypo-methylated states, suggesting complex regulation and differential involvement in the pan-heme landscape (**Figure 5**)

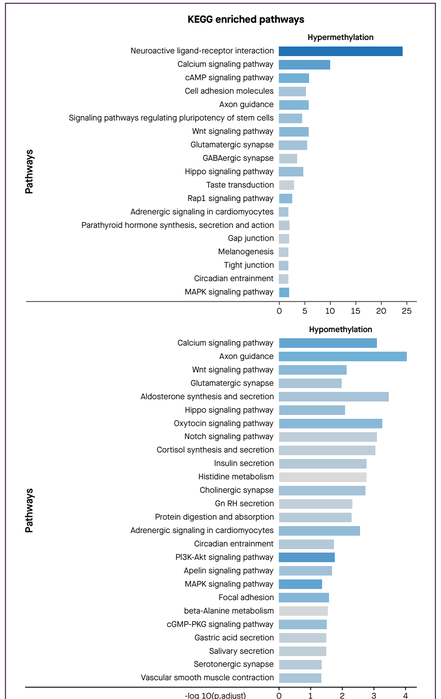
Figure 4. Pan-Heme DMRs Were Identified Between High- and Low-Risk Participant Groups.



Volcano plots of significant DMRs between high- and low-risk participant groups. Each point represents a significant DMR (q-value < 0.05 and delta coefficient > 0.1). The mean methylation difference between high- and low-risk participant groups is plotted along the x-axis. A deltaBeta > 0 represents hypermethylation and deltaBeta < 0 represents hypomethylation in the high-risk group.

DMR, differentially methylated region; heme, hematologic.

Figure 5. Epigenetic Signatures Were Identified Between High- and Low-Risk Participant Groups for Pan-Heme Neoplasms.



KEGG pathway enrichment of significant DMRs between high- and low-risk participants. The FDR-corrected q-value of pathway enrichment is reported for selected pathways.

DMR, differentially methylated region; FDR, false discovery rate; heme, hematologic; KEGG, Kyoto Encyclopedia of Genes and Genomes.

CONCLUSIONS

- We developed a novel, minimally-invasive, pan-heme prognostic classification model utilizing cfDNA targeted methylation sequencing technology that demonstrated proof-of-concept performance in representative heme neoplasm subtypes
- This cfDNA-based approach demonstrated potential for accurate risk stratification across a diverse range of heme neoplasms
- Although survival curve risk stratification was not significant for every heme subtype, this was likely due to small sample sizes. Due to the promising C-index results, follow-up studies will evaluate the methylation prognostic model in additional heme subtypes with larger sample sizes
- Our findings suggest that simple, minimally-invasive blood draw instead of a combination of multiple testing modalities offers a more efficient approach to disease prognosis across large heterogeneous groups of hematologic malignancies
- Future studies will evaluate the methylation prognostic model in another independent cohort and further assess the biological drivers of risk stratification in heme neoplasms

References

1. Scherer F et al. *Blood.* 2017; 130 (4): 440–452.
2. Taylor J et al. *Blood* 2017; 130 (4): 410–423.
3. Greenberg PL et al. *Blood* 2012; 120 (12): 2454–2465.
4. International Non-Hodgkin's Lymphoma Prognostic Factors Project. *N Engl J Med* 1993; 329:987-994.
5. Federico M et al. *J Clin Oncol.* 2009;27(27):4555-4562.
6. Hoster E et al. *Blood* (2008) 111 (2): 558–565.
7. International CLL-IPI working group. *Lancet Oncol.* 2016;17(6):779-790.
8. Liu MC et al. *Ann Oncol.* 2020;31(6):745-759.9.
9. Ishwaran, MS et al. Fast Unified Random Forests with randomForestSRC. 2023. https://www.randomforestsrc.org/articles/survival.html
10. Feng, H & Wu, H. *Quantitative Biology.* 2019:7(4), 327.
11. Kanehisa, M, et al. *Nucleic Acids Research.* 2017;45(D1):D353–D361.

Disclosures

Study funded by GRAIL, LLC. RS and YH are employees of GRAIL. AS, OV, and QL are employees of GRAIL with equity in the company and stock options in Illumina, Inc. OV has received research funding from GRAIL.

Acknowledgements

Funded by GRAIL, LLC. Editorial and graphic assistance provided by Prescott Medical Communications Group (Chicago, IL).