European Journal of Cancer 229 (2025) 115752

Contents lists available at ScienceDirect

European Journal of Cancer

journal homepage: www.ejcancer.com

Original research

AI-informed retinal biomarkers predict 10-year risk of onset of multiple hematological malignancies

Amritpal Singh [a], Ajay K. Nooka [b], Gourav Modanwal [c], Nieraj Jain [d], Madhav V. Dhodapkar [b], Sruthi Arepalli [d], Sagar Lonial [b], Anant Madabhushi [a,c,e,*]

[a] Emory University, Atlanta, USA
[b] Winship Cancer Institute, Emory University, Atlanta, USA
[c] Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA, USA
[d] Department of Ophthalmology, Emory University, Atlanta, USA
[e] Atlanta Veterans Administration Medical Center, Atlanta, USA

ARTICLE INFO

Keywords:
Hematological cancers
Leukemia
Myeloma
Lymphoma
Retinal biomarkers
Oculomics
Proteomics

ABSTRACT

Background: Early detection of hematological malignancies improves long-term survival but remains a critical challenge due to heterogeneity in clinical presentation. Chronic inflammation is a key driver in hematologic cancers and is known to induce compensatory microvascular changes. High-resolution, non-invasive retinal imaging can allow the quantification of microvascular changes for the early detection of hematological malignancies.

Methods: This study evaluated RetHemo, an explainable AI tool predicting hematological malignancy onset up to 10 years before diagnosis using retinal imaging in 1237 UK Biobank patients. Retinal vasculature features (curvature, tortuosity, branching angles) were extracted from segmented vessels, arteries, and veins, enabling high-risk subgroup identification and outperforming traditional clinical predictors.

Results: RetHemo demonstrated significant predictive performance for leukemia (c-index = 0.611, HR = 2.45, 95 % CI: 1.27–4.75, p = 0.027), myeloma (c-index = 0.636, HR = 6.69, 95 % CI: 2.06–21.65, p = 0.006). Unsupervised hierarchical clustering based on retinal vasculature features identified distinct high-risk subgroups for leukemia (p = 0.013), myeloma (p < 0.001), and lymphoma (p = 0.034). Serum proteomics analysis revealed significantly elevated levels of inflammatory proteins, including ITGAL and SLAMF7, in high-risk patients. Comparison with clinical variables showed that RetHemo outperformed traditional clinical and hematologic parameters in stratifying at-risk individuals.

Conclusion: These findings support the potential of AI-driven retinal biomarkers as a novel prognostic tool for early detection of hematological malignancies, enabling timely intervention and improved patient outcomes.

Statement of Significance

AI-driven retinal biomarkers derived from non-invasive fundus imaging predict the 10-year risk of leukemia, myeloma, and lymphoma, providing a novel approach for early detection and risk stratification of hematological malignancies.

1. Introduction

Hematological malignancies, including leukemia, myeloma, and lymphoma, remain challenging to detect early due to their variable clinical presentations and often subtle early symptoms. These cancers, among the most lethal, account for roughly 10 % of all malignancies, with mortality rates exceeding those of more common cancers like prostate and breast cancer. Despite advances in treatment, survival rates remain suboptimal, particularly in myeloma and leukemia. Diagnostic delays, often due to a combination of patient, systemic, and social factors, can worsen prognosis, with emergency diagnoses showing notably lower 3-year survival rates (40 %) compared to primary care diagnoses (77 %). [1,2] Thus, early identification and monitoring of high-risk individuals are crucial to enable timely intervention, prevent organ

* Correspondence to: 1750 Haygood Dr NE, N602, 6th floor, Health Sciences Research Building II (HSRB II), Emory University, Atlanta, GA 30306, USA.
E-mail address: anantm@emory.edu (A. Madabhushi).

https://doi.org/10.1016/j.ejca.2025.115752
Received 28 March 2025; Received in revised form 14 August 2025; Accepted 20 August 2025
Available online 30 August 2025

damage, and ultimately improve survival outcomes.

Chronic inflammation is a key driver in hematological cancer and is known to induce compensatory microvascular changes. [3,4] Increasing evidence suggests the presence of early ocular manifestations in hematological malignancies. Recent U.S. claims data (2011–2020) revealed that about 40 % of newly diagnosed multiple myeloma patients present with pre-existing ocular alterations. [5] Retinal vessel density changes have been reported in newly diagnosed acute leukemia, [6,7] remission of acute leukemia without overt retinopathy, [8] and newly diagnosed multiple myeloma patients using OCTA. [9] The presence of ocular manifestations at the time of diagnosis suggests that precursor stages of hematological cancer might manifest as microvascular remodeling in the retina.

Recent advances in AI and deep learning have enabled the detection

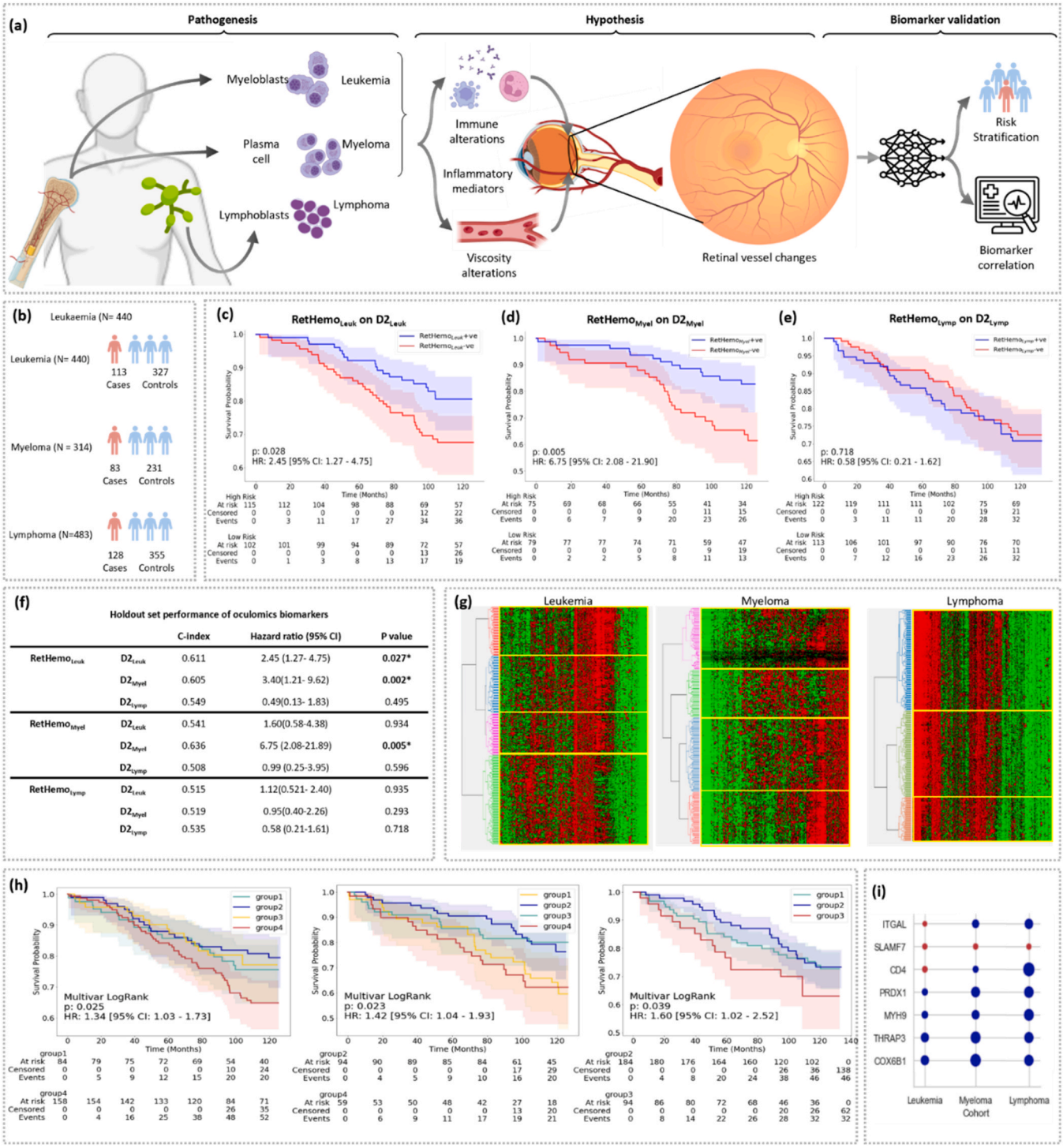

Fig. 1. (a) Conceptual diagram of the pathophysiology of retinal changes in hematological cancer. (b) Case(red) - control (blue) distribution for leukemia, myeloma, and lymphoma cohort (left to right) (c) KM curve for RetHemo$_{Leuk}$ (d) KM curve for RetHemo$_{Myel}$ (e) KM curve for RetHemo$_{Lymp}$ on (f) C-index and Hazard ratios for different models (g) Unsupervised clustering reveals oculomics based clustering of patients at risk of developing leukemia, myeloma, and lymphoma. (h) KM curves and multivariate long rank tests for groups based on unsupervised clustering (i) Differential expression in proteomics based on RetHemo groups (where size: p-value, red: significant).

and diagnosis of different systemic diseases from ocular imaging. [10–12] Color fundus imaging provides a non-invasive and accessible approach for visualizing and quantifying retinal microvascular morphology. In this context, we introduce RetHemo, an interpretable AI-driven retinal vasculature risk score designed to assess the likelihood of the onset of hematological malignancies for up to 10 years. To evaluate the association of RetHemo with these malignancies, we use multiple approaches, including supervised and unsupervised classification, followed by correlation analysis with proteomics to assess the biological basis of RetHemo predictions.

2. Methods

2.1. Study population

1370 hematological cancer patients were identified based on ICD 10 codes (Leukemia: lymphoid leukemia C91, myeloid leukemia C92; Myeloma: multiple myeloma C90; Lymphoma: diffuse non-Hodgkin C83, other non-Hodgkin lymphoma C85) from the UK Biobank cohort. To understand the risk of developing hematological cancer, only patients with pre-diagnosis fundus scans were selected. Since cardiovascular diseases are known to cause changes in retinal vasculature, patients with any major cardiovascular event before the retinal scan were excluded from the analysis. Controls were selected via 1:3 patient matching using patient age, sex, hypertension, and diabetes status. Psmpy version 0.3.13 (python) module was used for patient matching. The patient selection process is summarized in supplementary Figure S1.

Scans were randomly divided into 50–50 splits to create a train ($D1_{Leuk}$, $D1_{Myel}$, $D1_{Lymp}$) and a holdout set ($D2_{Leuk}$, $D2_{Myel}$, $D2_{Lymp}$). After patient matching, a total of 1237 patients were included in the analysis. (see Fig. 1) The leukemia cohort included 440 patients (D_{Leuk}), with 208 (53 %) males, the myeloma cohort (D_{Myel}) had 314 patients with 179 (57 %) males, and the lymphoma cohort (D_{Lymp}) had 440 patients with 279 (58 %) males. The average time to diagnosis was 4.89 years, with a maximum of 10.1 years. Table 1 describes the baseline patient demographics for case and matched control patients in each cohort.

For the leukemia cohort, patients were identified using the following ICD10 codes: lymphoid leukemia (C91, C910, C911, C912, C913, C914, C915, C916, C917, C918, C919) and myeloid leukemia (C92, C920, C921, C922, C923, C924, C925, C926, C927, C928, C929). Myeloma patients were selected using the ICD10 codes C900, C901, C902, and C903. For the lymphoma cohort, patients were identified using ICD10 codes for diffuse non-Hodgkin lymphoma (C83, C830, C831, C832, C833, C834, C835, C836, C837, C838, C839) and other non-Hodgkin lymphoma (C85, C850, C851, C852, C857, C859).

2.2. Feature extraction

The scans were randomly divided into two equal splits, with 50 % as the training set ($D1_{Leuk}$, $D1_{Myel}$, $D1_{Lymp}$) and the remaining 50 % as the holdout set ($D2_{Leuk}$, $D2_{Myel}$, $D2_{Lymp}$) for model training and final validation, respectively. Retinal scans were automatically segmented for vessels, arteries, and veins using a pre-trained convolutional neural network [25]. A Resnet50 convolutional neural network was trained to detect the fovea location and create concentric zones around the foveal center, allowing us to capture parafoveal microvasculature near the fovea avascular zone. We further segment the optic disk and optic cup using an ensemble of pre-trained convolutional neural networks to create zones around the optic disk to capture larger vessels near the optic disk. Two concentric zones were created around the optic disk, using optic disk diameter to capture larger vessels near the optic disk. Retinal vasculature features like curvature, tortuosity, and branching angle features were extracted separately for all types of vasculatures (vessel, artery, and vein) and with masked zone vasculature around the fovea and optic disk. Features were aggregated using summary statistics. The top three oculomics features were selected using only the train set. A Cox proportional regression model was trained on training sets ($D1_{Leuk}$, $D1_{Myel}$, $D1_{Lymp}$) to generate corresponding models ($RetHemo_{Leuk}$, $RetHemo_{Myel}$, $RetHemo_{Lymp}$) on the holdout sets ($D2_{Leuk}$, $D2_{Myel}$, $D2_{Lymp}$). C-index and hazard ratio were used to compare the predictive performance of each model.

2.3. Modeling

A Cox proportional regression model was trained using top three selected oculomics features on each respective training set ($D1_{Leuk}$, $D1_{Myel}$, $D1_{Lymp}$) to generate risk scores ($RetHemo_{Leuk}$, $RetHemo_{Myel}$, $RetHemo_{Lymp}$) that were subsequently validated on the corresponding holdout sets ($D2_{Leuk}$, $D2_{Myel}$, $D2_{Lymp}$). C-index and hazard ratio were used to compare the predictive performance of each model. We also performed an unsupervised hierarchical clustering analysis on vasculature features to identify distinct retinal phenotype clusters associated with unique risk profiles. Multivariate and univariate log-rank tests were performed to compare the risk status of clusters. Finally, we compared RetHemo predictions with antibody-based serum protein expression levels measured using the OLINK Proximity Extension Assay.

2.4. Model comparison

Clinical model $RetHemo_{clinical}$ (built using patient age, sex, hypertension, and diabetes status) and blood cell counts) and Blood cell based model $RetHemo_{Blood}$ (built using RBC, lymphocyte, monocyte, neutrophil, eosinophil, and basophil) at the time of imaging were trained and validated on respective training and validation sets.

Table 1
Baseline characteristics of patients included in three cohorts, with mean age of 65 years. Data are n (%) or median (range).

	Leukemia (D_{Leuk})		**Myeloma (D_{Myel})**		**Lymphoma (D_{Lymp})**	
Overall						
Count: N	440		314		483	
Age(yrs): mean(min/max)	65(45/78)		65(44/78)		65(44/77)	
Sex Male Female: N	208 (53 %)232 (47 %)		179 (57 %)135 (43 %)		279 (58 %)204 (42 %)	
Hypertension: N (%)	185 (42 %)		100 (32 %)		181 (37 %)	
Diabetes: N (%)	44 (10 %)		7 (2 %)		12 (3 %)	
Ethnicity (Black) N (%)	5 (1 %)		1 (<1 %)		4 (~1 %)	
Subgroups	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**
Count (N)	113	327	83	231	128	355
Age(yrs): mean(min/max)	66(45/78)	67(45/78)	66(45/78)	66(45/78)	66(44/75)	65(44/75)
Sex Male: N (%) Female: N (%)	56 (51 %)57 (49 %)	176 (54 %)151 (46 %)	46 (56 %)37 (44 %)	133 (57 %)98 (43 %)	70 (54 %)58 (45 %)	209 (59 %)146 (41 %)
Hypertension: N (%)	71 (37 %)	143 (44 %)	24 (29 %)	76 (32 %)	42 (33 %)	139 (39 %)
Diabetes: N (%)	10 (9 %)	34 (10 %)	4 (4 %)	4 (2 %)	3 (2 %)	10 (3 %)

2.5. Unsupervised clustering

For unsupervised hierarchical clustering, we used the clustergram package from MATLAB R2024b (RRID:SCR_001622), to find distinct retinal biomarker patterns associated with varying risk profiles in the entire cohort (i.e., D_{Leuk}, D_{Myel}, D_{Lymp}). First, the differential risk of developing different malignancies in all patient clusters was compared using the multivariate log-rank test. Subsequently, a univariate log-rank test was performed by merging subclusters with a prevalence higher than the mean prevalence as high-risk cohorts and others as low-risk cohorts. We report the hazard ratio along with its confidence interval and p-value of significance for both comparisons.

Proteomic analysis: The UK Biobank study encompasses 58,718 blood plasma samples with population-normalized measurements for over 2923 proteins, obtained using the antibody-based Olink Explore 3072 Proximity Extension Assay (PEA) platform. We identified eligible participants from our cohorts, i.e., 40 in D_{Leuk} (RetHemo+/- N = 20/19), 36 in D_{Myel} (RetHemo+/- N = 10/25), and 44 in D_{Lymp} (RetHemo+/- N = 15/27). For this study, we selected 22 protein markers previously demonstrated to have prognostic significance in leukemia and myeloma (Table S2) [13–15]. Missing values were removed, and no post-processing was done. To understand the differential expression of proteomic markers, we compared RetHemo +ve and -ve groups using a *t*-test. Given the relatively small sample size of patients with available proteomic data, we considered 0.1 or lower as the threshold to attain statistical significance.

Blood cell count: Hematological cancers often alter peripheral blood cell counts. Raised WBC has been shown to correlate with retinal changes in leukemic patients. [26] Increased WBC count has also been shown to be positively correlated with the presence of retinal infiltrates and overall survival rate [27]. We compare the mean values of blood cell counts amongst risk groups using the Wilcoxon rank test and compare blood biomarker trends between RetHemo groups to actual patient diagnosis.

2.6. Statistical analysis

The experimental design involved randomly splitting retinal fundus scans 50:50 into training and holdout validation sets: $D1_{Leuk}$, $D1_{Myel}$, $D1_{Lymp}$ for training, and $D2_{Leuk}$, $D2_{Myel}$, $D2_{Lymp}$ for validation. The study was retrospective, and blinding was not applicable; however, the AI model was trained without prior knowledge of clinical outcomes. No formal power calculation was performed, as the sample size was determined based on the available UK Biobank dataset, ensuring statistical power for subgroup analysis. Harrell's C-index was used to evaluate model discrimination, with hazard ratios (HR) and 95 % confidence intervals (CI) reported to quantify risk. Multivariate and univariate log-rank tests compared risk status among clusters, while Wilcoxon rank-sum tests were used for blood biomarker comparisons. In unsupervised clustering, multivariate log-rank tests assessed cluster risks. For proteomic analysis, T-tests compared RetHemo +ve and -ve groups, using a significance threshold of $p < 0.1$ due to the limited sample size. Wilcoxon rank-sum tests were also applied to compare WBC, RBC, and other hematological parameters between RetHemo risk groups and actual cancer diagnoses.

3. Results

3.1. RetHemo performance in hematological malignancies

We evaluated the utility of RetHemo for predicting the onset of hematological cancer in 10 years. $RetHemo_{Leuk}$ +ve individuals exhibited a significantly increased risk of developing leukemia, with a concordance index (c-index) of 0.611 and a hazard ratio (HR) of 2.45 (95 % CI: 1.27–4.75, p = 0.027). Notably, when controlling for clinicopathological variables, $RetHemo_{Leuk}$ remained an independent prognostic indicator for leukemia development (HR = 1.73, 95 % CI: 1.24–2.42, p = 0.028). On comparing clusters from unsupervised clustering, a difference in risk for developing leukemia was observed for the multivariate long-rank test (HR = 1.34, 95 % CI: 1.03–1.73, p = 0.025) and univariate log-rank test (HR = 1.60, 95 % CI: 1.10–2.33, p = 0.013). To understand expression of serum protein markers, we selected 22 protein markers previously demonstrated to have prognostic significance in leukemia and myeloma [13–15]. RetHemo+ leukemia patients had a differential expression for three proteins, including ITGAL (p = 0.01), a lymphocyte adhesion and migration protein, and SLAMF7 (p = 0.059), an immune modulation protein.

The $RetHemo_{Myel}$ +ve individuals had a significantly elevated risk of developing myeloma, with a c-index of 0.636 and an HR of 6.69 (95 % CI: 2.06–21.65, p = 0.006). Furthermore, when controlling for clinicopathological variables, $RetHemo_{Myel}$ was independently prognostic for the onset of myeloma (HR = 3.86, 95 % CI: 2.03–7.32, p = 0.002). On comparing clusters from unsupervised clustering, a difference in risk for onset of myeloma was observed via the multivariate (HR = 1.42, 95 % CI: 1.04–1.93, p = 0.023) and univariate log-rank tests (HR = 2.12, 95 % CI: 1.36–3.29, p < 0.001). On comparing serum protein levels, RetHemo+ myeloma patients had a differential expression of SLAMF7 (p = 0.059). These elevated levels of immune signaling proteins suggest the presence of inflammatory changes in high-risk myeloma patients, in turn manifesting as vasculature changes in the retina.

Univariate analysis of $RetHemo_{Lymp}$ yielded a c-index of 0.503 on $D2_{Lymp}$ and predicted high-risk cohort was not significantly higher (p = 0.868) risk for onset of lymphoma. On comparing clusters from unsupervised clustering, a difference in risk for onset of lymphoma was observed on multivariate (HR = 1.60, 95 % CI: 1.02–2.52, p = 0.039) and univariate log-rank test (HR = 1.54, 95 % CI: 1.03–2.30, p = 0.034). On comparing serum protein levels, RetHemo+ lymphoma patients had a differential expression of SLAMF7 (p = 0.068) and CD38, an immune mediator protein (p = 0.092), suggesting elevated levels of immune signaling in patients with retinal vasculature changes. Figs. 2–4 show results from hierarchical clustering. Prevalence of disease vs healthy group detected in each cluster and KM curves for leukemia, myeloma, and lymphoma cohorts, respectively. Figure S2 shows qualitative visualization for retinal vessel maps in peri-macular zones in $RetHemo_{Leuk}$ + and $RetHemo_{Myel}$ + cases.

3.2. Comparison with clinical variables

Clinical parameters (patient age, sex, hypertension, diabetes status) and blood cell counts (RBC, lymphocyte, monocyte, neutrophil, eosinophil, basophil) at the time of imaging were used to train a Cox regression model and develop $RetHemo_{clinical}$ and $RetHemo_{Blood}$, respectively. Neither $RetHemo_{Clinical}$ ($D2_{Leuk}$: p = 0.54, $D2_{Myel}$: p = 0.55, $D2_{Lymp}$: p = 0.25) nor $RetHemo_{Blood}$ ($D2_{Leuk}$: p = 0.41, $D2_{Myel}$: p = 0.18, $D2_{Lymp}$: p = 0.28) significantly differentiated between high and low-risk groups across the three cohorts. Table S1 shows the model performance for all three cohorts. Univariate analysis of $RetHemo_{Leuk}$ on $D2_{Myel}$ had a c-index of 0.605, and the individuals identified as high-risk had a significantly higher (p = 0.002) risk for developing myeloma (HR 3.40, CI 95 %: 1.21–9.62). Fig 5 shows the KM curves for different combinations. Univariate analysis of $RetHemo_{Myel}$ on $D2_{Leuk}$ had a c-index of 0.541, but the predicted high-risk cohort was not significantly higher (p = 0.934) risk for developing leukemia. This suggests that retinal changes in myeloma from increased abnormal protein are unique and not found in leukemia [16]. However, leukemia-induced features are found in both, possibly due to a common inflammatory pathway.

3.3. RetHemo correlation with proteomics and blood cells

In the Leukemia cohort, we see that RetHemo score-based groups showed statistically significant differences in lymphocyte, neutrophil, and neutrophil-to-lymphocyte ratio. $RetHemo_{Leuk}$ +ve individuals had a

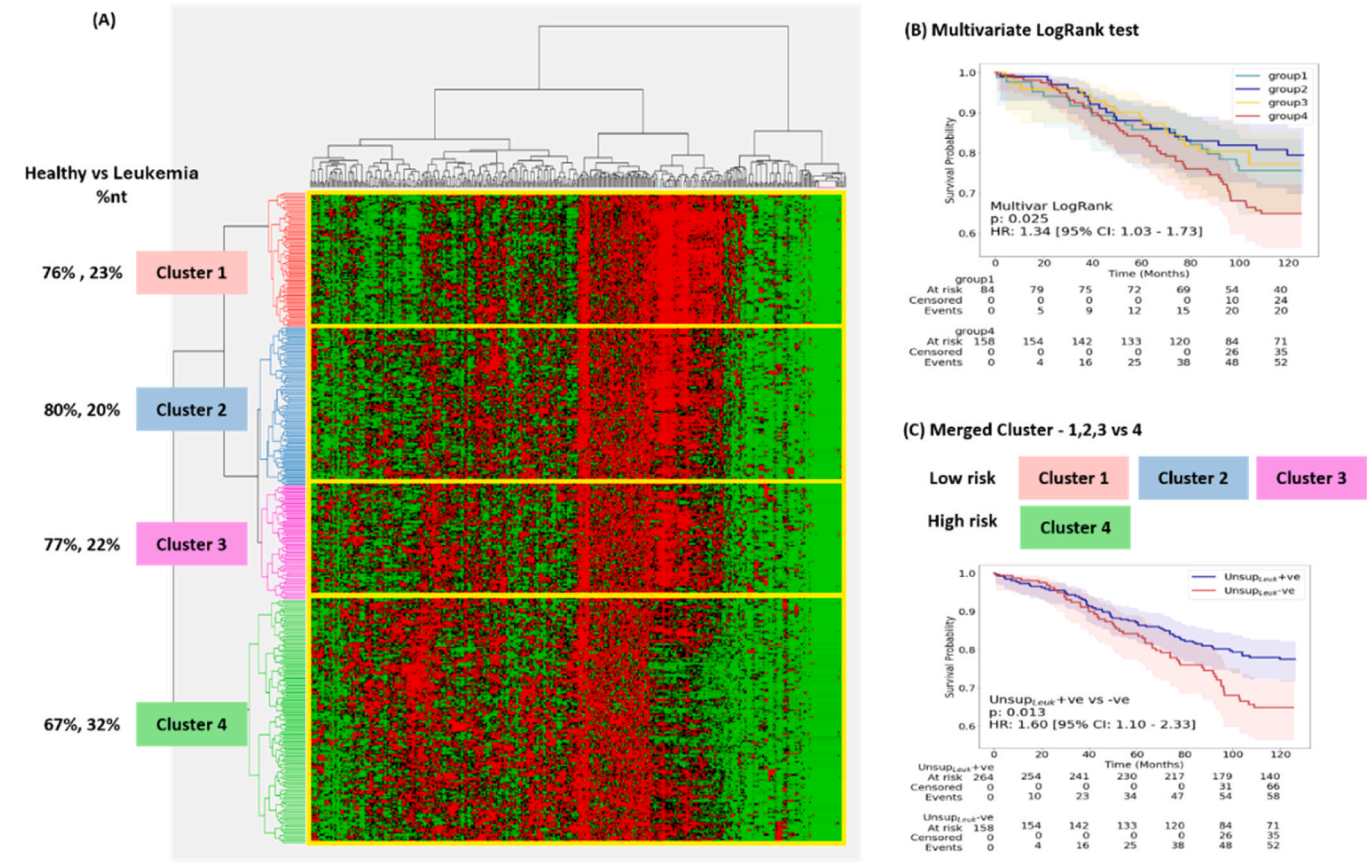

Fig. 2. Unsupervised hierarchical clustering based on retinal vasculature features on leukemia D_{Leuk} cohort. (A) Each row represents a patient, and each column represents a retinal feature. (B) multivariate and (C) univariate log-rank analysis with KM curves showing significant differences (p = 0.025 and 0.013 respectively) in the likelihood of developing leukemia.

significantly higher (p = 0.024) lymphocyte count of 0.281 (CI 95 %: 0.248–0.309) compared to the low group, 0.228 (CI 95 %: 0.209–0.252). $RetHemo_{Leuk}$ +ve individuals had a significantly lower (p = 0.022) neutrophil count of 0.666 (CI 95 %: 0.637–0.697) compared to the low group, 0.727 (CI 95 %: 0.637–0.697). Low neutrophil counts in leukemia have a poor prognosis. [17]. WBC count levels can be predicted using fundus imaging [18]. An increase in lymphocyte count might increase leukemic infiltration, which is a poor prognostic factor in leukemia. On univariate analysis of Whole blood viscosity (WBV) parameters, $RetHemo_{Myel}$ +ve patients had elevated high shear rate (p = 0.063) and low shear rate (p = 0.071), with no significant differences in $RetHemo_{Leuk}$ +ve and $RetHemo_{Lymp}$ +ve.

$RetHemo_{Myel}$ +ve individuals had a significantly lower (p = 0.013) monocyte count of 0.342, CI 95 %: [0.307–0.378] compared to the low group 0.411, CI 95 %: [0.369–0.448]. Fig. 6.a and b illustrate the final p-values across the three cohorts and the volcano plots for the analysis, respectively. Interestingly, expression trends in $RetHemo_{Myel}$+ groups closely align with protein expression patterns observed based on actual myeloma diagnoses of patients. Multiple myeloma has a significantly lower classical monocyte count. [19], higher non-classical monocytes [20], and was of prognostic value. While multiple myeloma normally results in elevated levels of monocytes at diagnosis, 82 % of multiple myeloma patients have monoclonal gammopathy of undetermined significance (MGUS) 8 years before multiple myeloma diagnosis. [16]. Trafficking of immune cells to the BM and localization to the tumor site is one of the first critical steps to shape the multiple myeloma microenvironment and may affect disease development and progression. [21]. Cytokines from tumors attract and recruit circulating monocytes to generate TAMs at the tumor site. [22]. TAMs promote plasma cell proliferation and angiogenesis, which supports MM progression [3]. We hypothesize that this monocyte trafficking to the tumor site is a potential reason for the reduced monocyte count found in this high-risk group.

3.4. Classification and time to event prediction

To understand the discriminative nature of retinal features, we train a classification model to predict if a patient would develop hematological cancer. Similar to Cox models, we use training sets ($D1_{Leuk}$, $D1_{Myel}$, $D1_{Lymp}$), followed by validation on the holdout sets ($D2_{Leuk}$, $D2_{Myel}$, $D2_{Lymp}$). On the holdout set $D2_{Leuk}$, the Xgboost classifier model demonstrated an AUC of 0.595 (95 % CI: 0.525–0.696) in detecting leukemia, with a sensitivity of 0.452 (CI 95 % 0.318–0.569) and a specificity of 0.699 (CI 95 % 0.623–0.768). To assess the predictive value of the classification signature, a kaplan-meier analysis was performed using the machine learning-derived probabilities as a risk score. On univariate analysis, high-risk patients on D2Leuk had a significantly elevated risk of developing leukemia (HR 3.16, 95 % CI: 1.42–7.06; p = 0.029) with a c-index of 0.582. Although trained solely for classification, the prediction probabilities effectively identified an increased risk of leukemia onset. Conversely, the Xgboost models for myeloma and lymphoma failed to prognosticate disease onset, with AUC values of 0.518 (95 % CI: 0.427–0.618) for myeloma and 0.539 (95 % CI: 0.461–0.620) for lymphoma.

4. Discussion

This study introduces RetHemo, an AI-driven retinal vasculature biomarker for predicting the 10-year risk of hematological malignancies based on fundus imaging. Leveraging non-invasive fundus images from the UK Biobank, RetHemo demonstrated significant prognostic

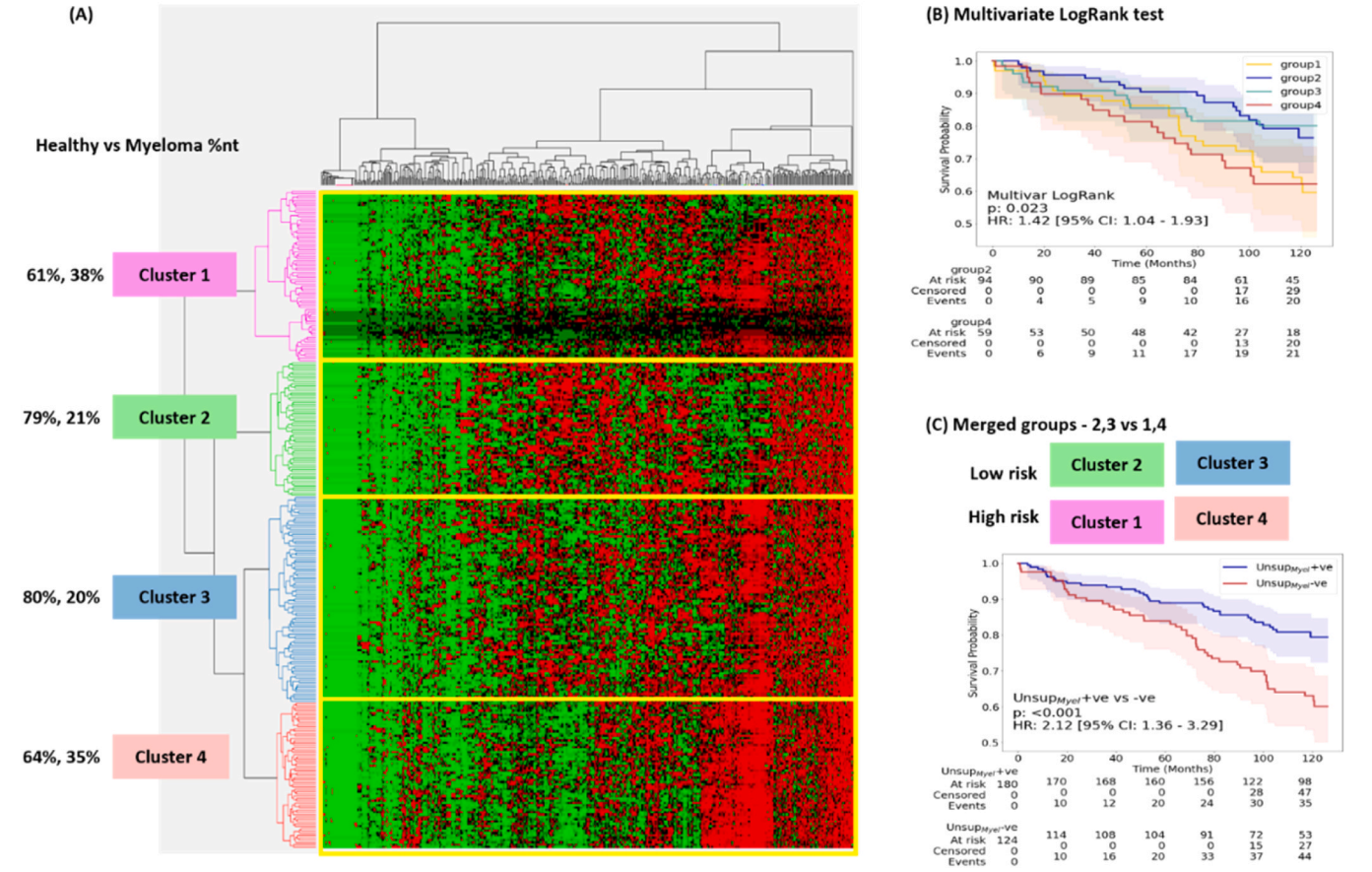

Fig. 3. Unsupervised hierarchical clustering based on retinal vasculature features on myeloma D_{Myel} cohort. (A) Each row represents a patient, and each column represents a retinal vasculature feature. (B) multivariate and (C) univariate log-rank analysis with KM curves showing significant differences ($p = 0.023$ and <0.001 respectively) in the likelihood of developing myeloma.

capabilities for leukemia, myeloma, and lymphoma. We compare the regions of signature across different hematological cancers and cross-test signatures between diseases. Finally, we explore the correlation of RetHemo with clinical parameters and serum protein levels to understand its biological underpinning in disease development. By enabling early detection and insights into immune-tumor interactions, RetHemo has the potential to advance prognostic strategies for hematological malignancies.

RetHemo shows promising results in differentiating patients with a high risk of developing leukemia and myeloma. RetHemo +ve patients had a two-fold higher risk of developing leukemia and six times higher risk of developing myeloma. To reduce the bias of feature selection, we repeat the experiment using unsupervised clustering. Clusters generated using retinal vascular features had a differential risk of developing leukemia, myeloma, and lymphoma on multivariate analysis. This signature stays independently prognostic of risk after adjusting for clinical variables. The non-invasive nature of fundus imaging for retinal vasculature quantification makes it an appropriate modality for detecting vascular changes and risk assessment in hematological cancer.

A crucial finding is characterizing the domain of the affected vasculature. In Leukemia, an increased artery tortuosity and vessel length in peri-macular regions was observed. These findings align with existing literature, [6–8] reporting retinal manifestations like dilatation and tortuosity of retinal vessels due to vascular stasis, occlusion, and retinal ischemia in leukemia at the time of diagnosis. Retinal vessel density changes in radial peripapillary capillaries ($p < 0.001$) have been reported in newly diagnosed acute leukemia, remission acute leukemia without overt retinopathy, and newly diagnosed multiple myeloma cases using OCTA. Chronic inflammation is a key driver in hematologic cancers, inducing the state of ischemia and hypoxia followed by a compensatory increase to maintain the blood supply to the retina. [23] Hypoxic-ischemia can induce the expression of hypoxia-inducing factor-1α and its target genes, such as vascular endothelial growth factor (VEGF), which disrupts the blood-retinal barrier and leads to retinal edema and microvascular changes. Bone marrow alteration in blood cell production can increase the risk of clotting, leading to blockages and fluid leakage in the micro-vessels of the retina, reducing blood flow and oxygen supply. RetHemo+ leukemia patients exhibited distinct expression patterns for key immune-related proteins, including ITGAL, involved in lymphocyte adhesion and migration, and SLAMF7, which plays a role in immune modulation. These proteomic insights suggest that retinal vascular changes in leukemia may be linked to immune or inflammatory pathways. However, further research is needed to clarify the causal mechanisms underlying these observations.

In contrast to leukemia, the myeloma signature comprises vein length and artery branching angle changes in peri-macular regions. Myeloma pathogenesis can sometimes start a few decades prior to the diagnosis, with an increase in abnormal protein or paraproteinemia during precursor stages. [16] Viscosity is a measure of blood's internal resistance to flow. Increased viscosity can result either from a deformity in the shape of the red blood cells or from a pathological elevation of any of the protein levels, and can cause reduced blood flow, tissue hypoperfusion, and lead to vascular leakage, and subsequent static hypoxia-induced microvascular changes. [24]

Interestingly the unique signatures identified for the leukemia, myeloma and were not prognostic when applied to different cancers, suggesting the presence of unique prognostic phenotypes, specific to the individual hematological cancers. This suggests that retinal changes in

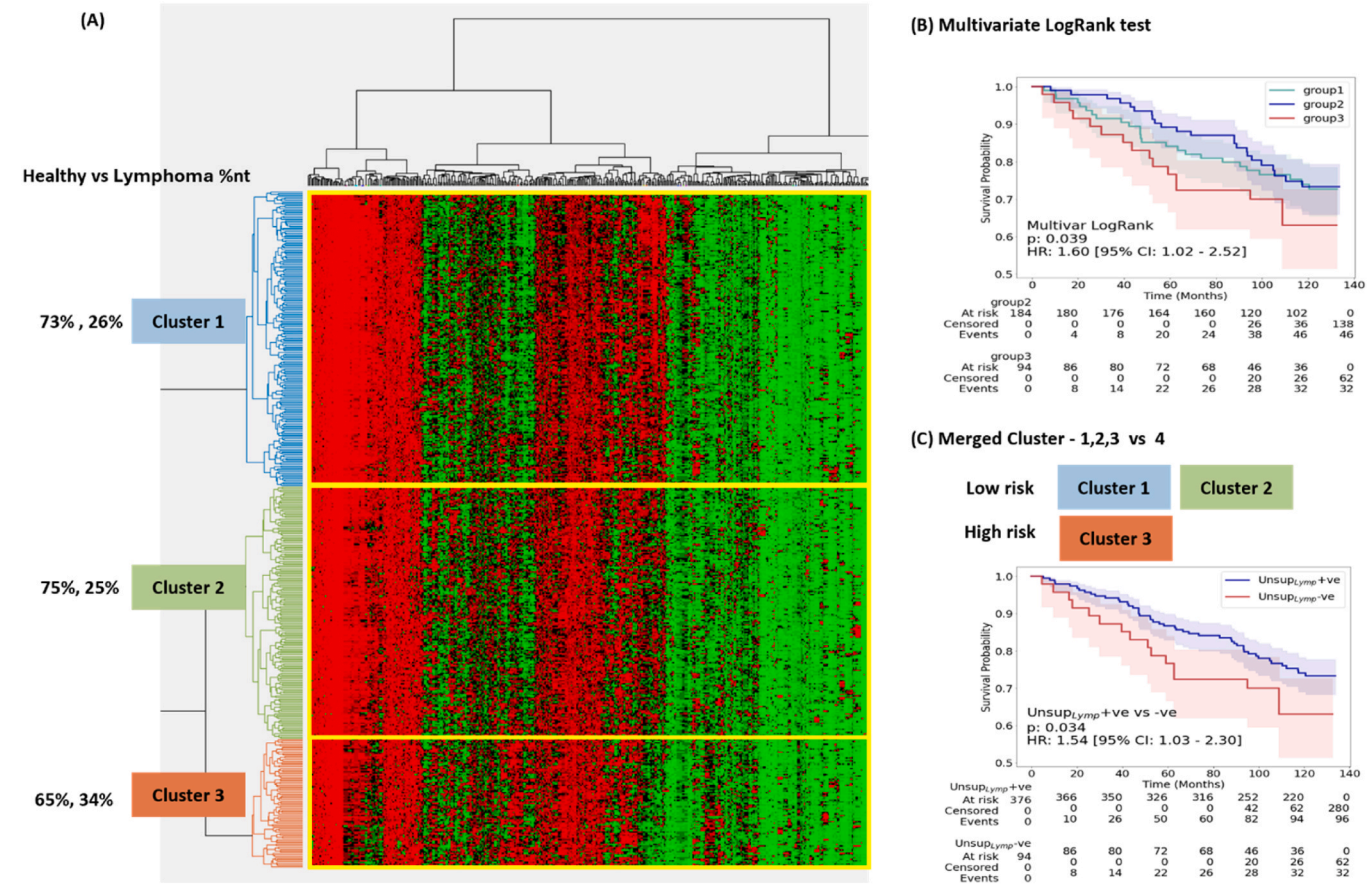

Fig. 4. Unsupervised Hierarchical clustering based on retinal vasculature features on lymphoma D_{Lymp} cohort. (A) Each row represents a patient, and each column represents a retinal vasculature feature. (B) multivariate and (C) univariate log-rank analysis with KM curves showing significant differences ($p = 0.039$ and 0.034 respectively) in the likelihood of developing lymphoma.

myeloma from increased abnormal protein are unique and not found in leukemia. [16] However, leukemia-induced features are found in both, possibly due to a common inflammatory pathway.

We acknowledge that our study did have its limitations. This study was restricted to adult patients from a single-site cohort. Efforts are underway to validate this in a multi-institutional setting. Unlike leukemia and myeloma, which are characterized by widespread bone marrow infiltration, marrow involvement in lymphoma is less common and localized to lymph nodes, potentially explaining the lack of a prognostic signal in lymphoma patients. This cohort was a predominately white population. Further prospective validation across multiple sites and in a diversity of ethnicities is needed for a more comprehensive understanding of RetHemo's utility.

We acknowledge prior studies demonstrating the role of retinal vessels in stroke, cardiovascular disease, and chronic kidney disease. Our findings highlight the complementary role of retinal vasculature in hematological cancers, suggesting the potential to extend existing retinal biomarkers into broader cardio-oncological investigations. Despite the aforementioned limitations, this study presents a novel application of AI and retinal imaging for predicting hematological cancer risk, providing a promising non-invasive and accessible approach to complement traditional blood-based and clinical risk assessments. By integrating oculomics into cancer risk evaluation, clinicians could enhance early detection strategies, especially in asymptomatic individuals, enabling early treatment initiation, preventing disease progression, and improving outcomes. Furthermore, RetHemo's correlation with immune-modulatory markers highlights its promise in illuminating cancer's systemic effects, reinforcing its biological relevance in hematological malignancies. Understanding these changes can help better understand biological pathways associated with microvascular changes in hematological cancers, potentially enabling the development of new drug lines. This approach could, in turn, enable large-scale, routine screening for early hematological cancer detection and foster deeper exploration into the systemic and retinal manifestations of oncogenic processes.

Code availability

The source code is available at GitHub: https://github.com/Amritpal-001/RetHemo

CRediT authorship contribution statement

Madhav V. Dhodapkar: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Conceptualization. **Nieraj Jain:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Conceptualization. **Anant Madabhushi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Sagar Lonial:** Writing – review & editing, Writing – original draft, Visualization, Validation. **Sruthi Arepalli:** Writing – review & editing, Writing – original draft, Visualization, Validation, Conceptualization. **Ajay K. Nooka:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **singh Amritpal:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Gourav Modanwal:** Writing – review & editing, Writing – original draft, Visualization, Validation,

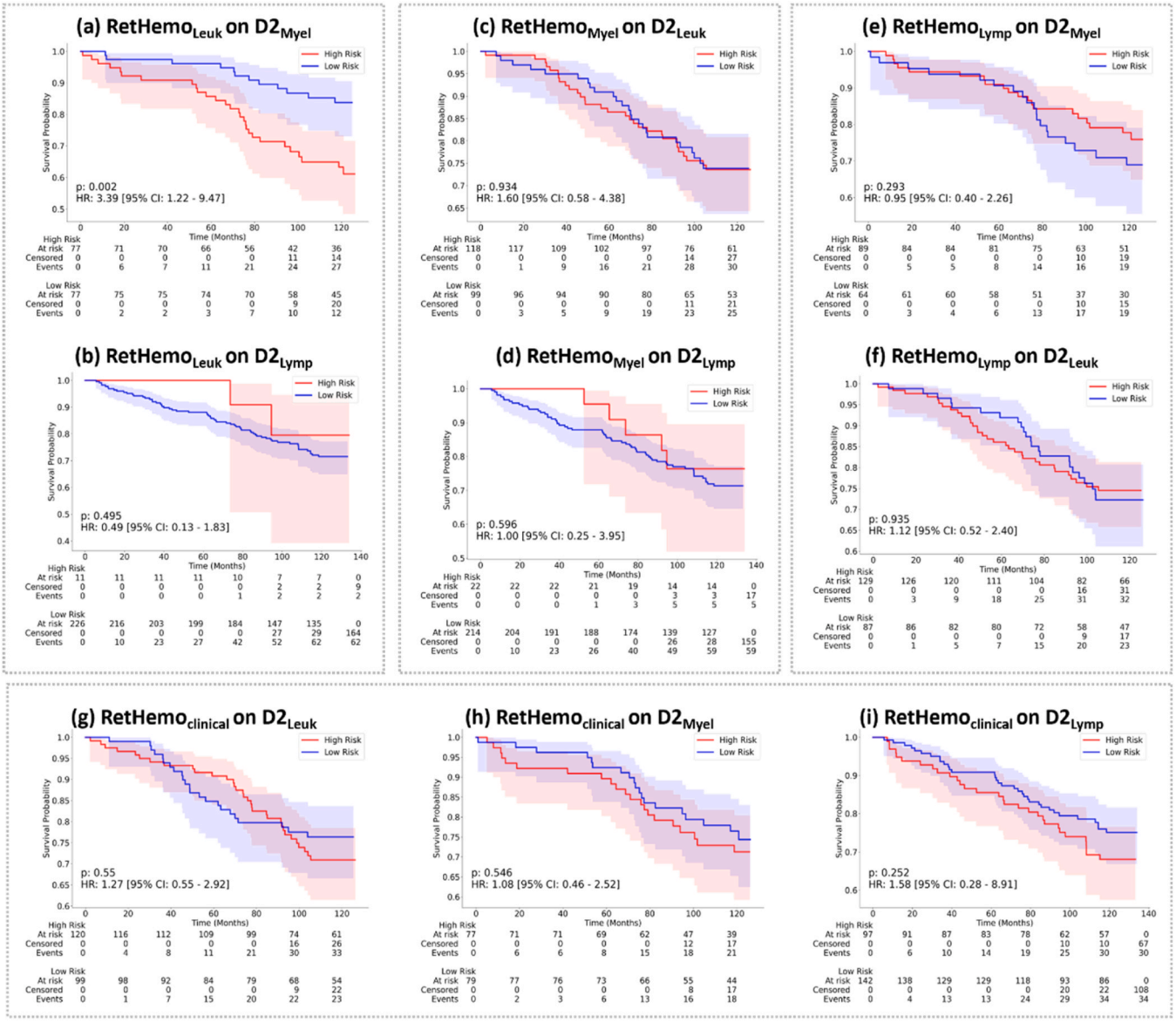

Fig. 5. KM curve analysis showing cross-testing of (a-b) RetHemo$_{Leuk}$ on D2$_{Myel}$, D2$_{Lymp}$ (c-d) RetHemo$_{Myel}$ on D2$_{Leuk}$, D2$_{Lymp}$ (e-f) RetHemo$_{Lymp}$ on D2$_{Leuk}$, D2$_{Myel}$ and (g-i) RetHemo$_{Clinical}$ on D2$_{Leuk}$, D2$_{Myel}$ and D2$_{Lymp}$.

Supervision, Methodology, Investigation, Data curation, Conceptualization.

Ethics approval and consent to participate

The UK Biobank dataset was ethically obtained from UK Biobank (application number 72280). Additionally, UK Biobank has received Research Tissue Bank approval from the North West Multi-Centre Research Ethics Committee and holds a Human Tissue Authority license.

Funding

Research reported in this publication was supported by the National Cancer Institute under award numbers R01CA249992-01A1, R01CA216579-01A1, R01CA257612-01A1, R01CA264017-01, R01CA268287-01A1, U01CA113913-16A1 U01CA239055-01, U01CA269181-01, U24CA274494-01 and U54CA254566-01.

Declaration of Competing Interest

Anant Madabhushi is an equity holder in Picture Health, Elucid Bioimaging, and Inspirata Inc. Currently, he serves on the advisory board of Picture Health, and SimBioSys. He currently consults for Takeda Inc. He also has sponsored research agreements with AstraZeneca and Bristol Myers-Squibb. His technology has been licensed to Picture Health and Elucid Bioimaging. He is also involved in 2 different R01 grants with Inspirata Inc. He also serves as a member of the Frederick National Laboratory Advisory Committee. Ajay K. Nooka has served on advisory boards and received honorarium from Adaptive Biotechnologies, Amgen, Astrazeneca, Bristol Myers Squibb, Cellectar biosciences, GlaxoSmithKline, Janssen, K36 therapeutics, ONK therapeutics, Pfizer, Sanofi, Sebia, and Takeda; received grant/research support (to university) from Aduro Biotech, Amgen, Arch Oncology, Bristol Myers Squibb, Cellectis, Genentech, GlaxoSmithKline, Janssen, Karyopharm, Kite Pharma, Merck, Pfizer, and Takeda; and received grant/research support for investigator-initiated studies from Amgen, GlaxoSmithKline, Janssen, Merck, and Takeda. Madhav V. Dhodapkar has served on advisory boards for Lava Therapeutics, BMS, Janssen, and

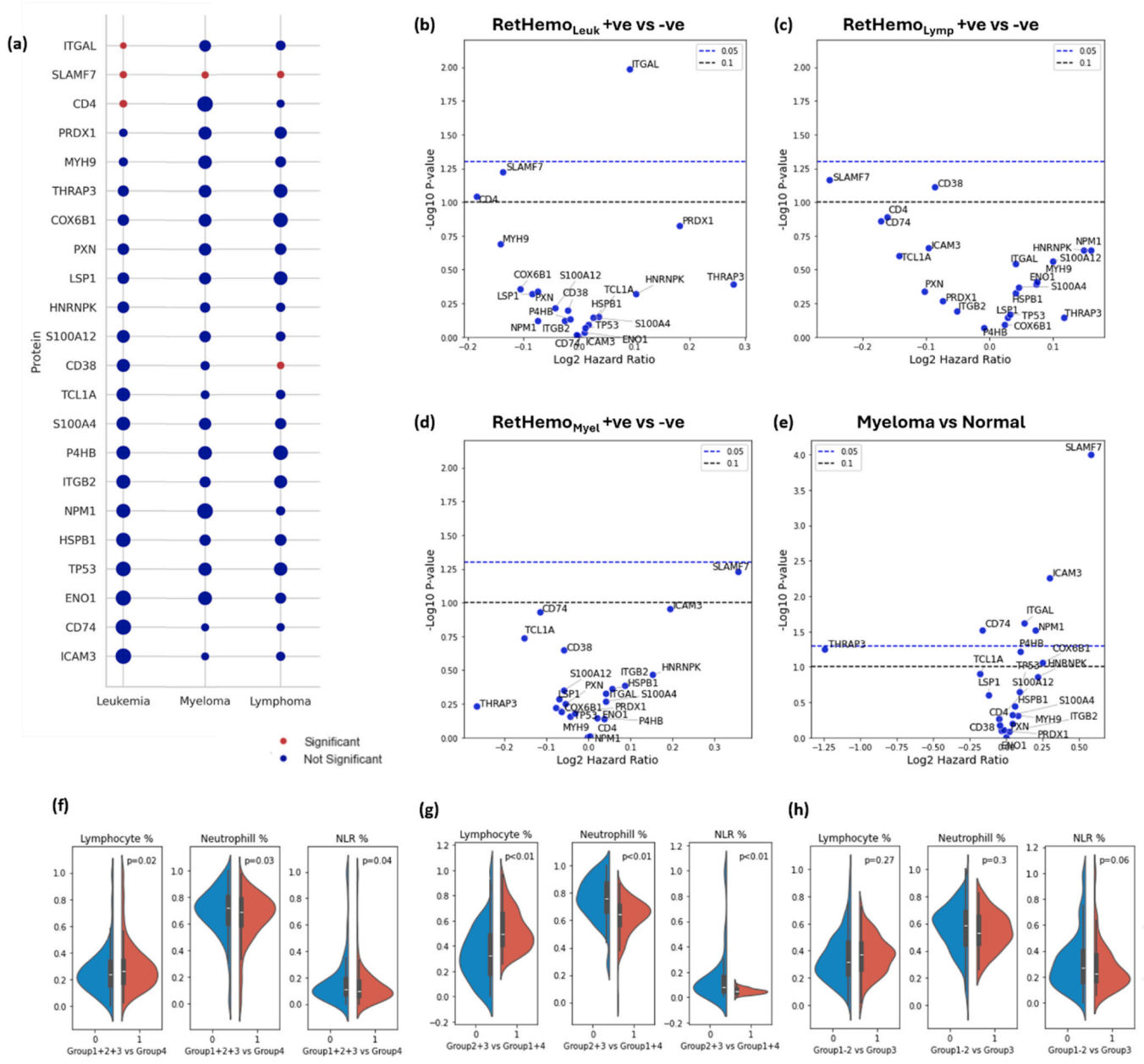

Fig. 6. (a) Differential expression of Proteomics markers based on RetHemo status (+ vs – groups). Volcano plots show differential protein expression in (b) leukemia using RetHemo$_{Leuk}$, (c) lymphoma using RetHemo$_{Lymp}$, (d) myeloma using RetHemo$_{Myel}$. (e) myeloma using diagnosis status. The volcano plot for proteins in myeloma using RetHemo$_{Myel}$ in (c) follows a similar trend as patient myeloma diagnosis in (e). Panels (f-h) show violin plots comparing lymphocyte, neutrophil, and neutrophil to lymphocyte ratio (NLR) using unsupervised clusters for leukemia, myeloma, and lymphoma, respectively.

Sanofi. Sagar Lonial has served in a consulting or advisory role for AbbVie, Amgen, Bristol-Myers Squibb, Celgene, GlaxoSmithKline, Janssen Oncology, Juno Therapeutics, Merck Novartis, and Takeda. Sruthi Arepalli has served as a consultant for AbbVie, and Genentech. Amritpal Singh, Gourav Modanwal, and Nieraj Jain have no competing interests.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejca.2025.115752.

Data availability

The UK Biobank test dataset was acquired through UK Biobank (application number 72280). Due to patient privacy concerns and the lack of informed consent for data sharing, the dataset cannot be publicly shared. However, researchers are encouraged to request access directly via the UK Biobank data portal.

References

[1] Blood Cancer UK. (2019) Report: The Current State of Blood Cancer Diagnosis in England.
[2] Abel GA, et al. Delays in referral and diagnosis for chronic hematologic malignancies: a literature review. Leuk Lymphoma 2008;49:1352–9.
[3] Schaefer C, et al. Time-course-dependent microvascular alterations in a model of myeloid leukemia in vivo. Leukemia 2008;22:59–65.
[4] Passaro D, et al. Increased vascular permeability in the bone marrow microenvironment contributes to disease progression and drug response in acute myeloid leukemia. Cancer Cell 2017;32:324–341.e6.
[5] Ailawadhi S, et al. Prevalence of ocular comorbidities in patients with multiple myeloma: an analysis of U.S. Claims data. 8038–8038 JCO 2022;40. 8038–8038.

[6] Zhou C, et al. Increased peripapillary capillaries in patients with acute leukemia by using optical coherence tomography angiography. Photo Photodyn Ther 2023;42:103569.
[7] Cicinelli MV, et al. Retinal microvascular changes in patients with acute leukemia. Retina 2022;42:1762–71.
[8] Zhou M, et al. Detection of retinal microvascular changes with optical coherence tomography angiography in patients with acute leukemia without retinopathy. Ophthalmol Ther 2024;13:1145–57.
[9] Dursun ME, et al. Retinal microvascular changes in patients with multiple myeloma: a study based on optical coherence tomography angiography. Curr Eye Res 2022;47:874–81.
[10] Zhang J, et al. RETFound-enhanced community-based fundus disease screening: real-world evidence and decision curve analysis. npj Digit Med 2024;7:1–9.
[11] Zekavat SM, et al. Phenome- and genome-wide analyses of retinal optical coherence tomography images identify links between ocular and systemic health. Sci Transl Med 2024;16:eadg4517.
[12] Ortín Vela S, et al. Phenotypic and genetic characteristics of retinal vascular parameters and their association with diseases. Nat Commun 2024;15:9593.
[13] Setayesh SM, et al. Targeted single-cell proteomic analysis identifies new liquid biopsy biomarkers associated with multiple myeloma. npj Precis Onc 2023;7:1–10.
[14] Griffen TL, et al. Proteomic profiling based classification of CLL provides prognostication for modern therapy and identifies novel therapeutic targets. Blood Cancer J 2022;12:1–14.
[15] Alsagaby SA, Alhumaydhi FA. Proteomics insights into the pathology and prognosis of chronic lymphocytic leukemia. Saudi Med J 2019;40:317.
[16] Fairfield H, Falank C, Avery L, Reagan MR. Multiple myeloma in the marrow: pathogenesis and treatments. Ann N Y Acad Sci 2016;1364:32–51.
[17] Chen X, et al. Low absolute neutrophil count during induction therapy is an adverse prognostic factor in childhood acute lymphoblastic leukaemia. Ann Hematol 2021;100:2269–77.
[18] Rim TH, et al. Prediction of systemic biomarkers from retinal photographs: development and validation of deep-learning algorithms. Lancet Digit Health 2020;2:e526–36.
[19] Zahran AM, et al. Higher proportion of non-classical and intermediate monocytes in newly diagnosed multiple myeloma patients in Egypt: a possible prognostic marker. Afr J Lab Med 2021;10:129.
[20] Edwards CV, Yildirim C, Fillmore N, Munshi NC. Monocytosis at diagnosis is associated with decreased overall survival in patients with newly diagnosed multiple myeloma. Blood 2020;136:8–9.
[21] Beider K, et al. Multiple myeloma cells recruit tumor-supportive macrophages through the CXCR4/CXCL12 axis and promote their polarization toward the M2 phenotype. Oncotarget 2014;5:11283–96.
[22] Sun J, et al. Tumor-associated macrophages in multiple myeloma: advances in biology and therapy. J Immunother Cancer 2022;10:e003975.
[23] Hasselbalch HC, et al. Myeloproliferative blood cancers as a human neuroinflammation model for development of Alzheimer's disease: evidences and perspectives. J Neuroinflamm 2020;17:248.
[24] Ferreira Santos da Cruz N, Milhomens Filho JAP, Ferraro DMN, Polizelli MU, de Moraes Ambrogini NSB. Hyperviscosity retinopathy and immunogammopathy maculopahy as new onset of multiple myeloma. Case Rep Ophthalmol 2021;12:578–84.
[25] Galdran A, et al. State-of-the-art retinal vessel segmentation with minimalistic models. Sci Rep 2022;12:6174.
[26] Guyer DR, et al. Leukemic retinopathy: relationship between fundus lesions and hematologic parameters at diagnosis. Ophthalmology 1989;96:860–4.
[27] Abu El-Asrar AM, Al-Momen A-K, Kangave D, Harakati MS, Ajarim DS. Correlation of fundus lesions and hematologic findings in leukemic retinopathy. Sage 1996. https://doi.org/10.1177/112067219600600213?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed.