Blood 146 (2025) 5756–5757

The 67th ASH Annual Meeting Abstracts

POSTER

653. MULTIPLE MYELOMA: CLINICAL AND EPIDEMIOLOGICAL

Machine learning uncovers prognostically distinct myeloma cast nephropathy phenotypes not captured by standard risk stratification systems

Michael Hughes[1], *Yiran Peng*[1], *Shikun Wang*[2], *Andrew Laine*[3], *Jai Radhakrishnan*[4], *Metodi Balev*[5], *Divaya Bhutani*[1], *Markus Mapara*[1], *Andrew Eisenberger*[1], *Suzanne Lentzsch*[1], *Rajshekhar Chakraborty*[1]

[1] Columbia University Irving Medical Center, Hematology/Oncology, New York, United States
[2] Columbia University Irving Medical Center, Biostatistics, New York, United States
[3] Columbia University School of Engineering, Biomedical Engineering, New York, United States
[4] Columbia University Irving Medical Center, Nephrology, New York, United States
[5] Johns Hopkins Medicine, Oncology, Baltimore, United States

Abstract **BACKGROUND** Myeloma cast nephropathy (MCN), a driver of renal failure in 10-15% of newly diagnosed multiple myeloma (NDMM), remains understudied. Current risk classification systems, including the 2025 IMS-IMWG high-risk disease definition [PMID 40489728], may not adequately stratify MCN patients in part due to reduced renal clearance of β-2-microglobulin (B2M). Unsupervised machine learning (ML) workflows have detected previously unrecognized data structures in other fields. We applied ML to a real-world NDMM cohort to explore latent MCN phenotypes.

METHODS We describe 332 adult patients (59 MCN, 273 contemporary controls) seen 5/2016-11/2024. MCN was defined per prior criteria [PMID 39989117]. We trained 2 feed-forward autoencoders on the overall and MCN-specific cohorts to generate compressed latent representations. Input features included free light chain (FLC) titers and urine protein levels over 6 months. We excluded outcome-related features. We also excluded baseline eGFR and creatinine in the full cohort. We reduced data dimensionality with uniform manifold approximation and projection (UMAP) or principal component analysis (PCA) and performed clustering on the resulting embeddings with Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) or *k*-means. We analyzed overall survival (OS) with Kaplan-Meier and Cox techniques. To prevent disproportionate exclusion of MCN patients from phenotype discovery, we discarded clustering attempts in which >20% of any noise labels were assigned to MCN patients. We excluded data labeled by clustering as "noise" from survival analysis. We derived cluster-based OS groups, performed feature comparisons, and identified noncollinear drivers using pooled Shapley additive explanations.

RESULTS Median ISS and R-ISS were both 2 (IQR 1-2), but MCN patients had higher stages ($p<0.001$). Median number of high-risk cytogenetic abnormalities ["Kaiser HRCA", PMID 39965171] was 1 (IQR 0-1). At diagnosis, MCN patients had significantly lower eGFR (20.4mL/min/1.7m^2 vs 72.3, $p<0.001$) and hemoglobin (9.1g/dL vs 10.2, $p<0.001$); and higher B2M (14.2mg/dL vs 5.5, $p<0.001$) and urine protein (3.300g/d vs 0.336, $p<0.001$). Median time to treatment was shorter in MCN (6 vs 17 days, $p<0.001$). ASCT rates did not differ. Six-month overall response rate favored MCN (94.5% vs 91%, $p=0.058$). There were 49 deaths (38 controls, 11 MCN). Median follow-up was 33.6 months (29.6-36.5). Median OS was not reached; 12- and 18-month OS were 95.3% and 92.8%. In univariate Cox models, ISS stages II/III, Kaiser HRCA=1, and age were the largest significant OS predictors. In multivariable models, ISS III (HR 8.13, $p=0.008$) remained significant. Kaiser HRCA=1 (HR 2.19, $p=0.077$) and ISS II (HR 4.07, $p=0.062$) showed borderline associations.

UMAP/HDBSCAN on the full cohort identified 6 clusters and 1 noise cluster (n=78, 23.5%): A (n=15, 4.5%); B (65, 19.6%); C (49, 14.8%); D (42, 12.7%); E (68, 20.5%); and F (15, 4.5%). OS differed significantly across clusters ($p=0.004$); clusters C and F were MCN-enriched (49.0% and 73.3% respectively; $p<0.001$). We categorized OS risk across clusters:

- Low risk (A+B+F, reference)
- Moderate risk (D+E, HR 3.57, $p=0.020$)
- High risk (C, HR 7.86, $p<0.001$).

The "high risk" group was enriched for high-stage ISS and R-ISS. Dominant cluster drivers included baseline FLC, lambda FLC, M protein, hemoglobin, and IgA. In the MCN subgroup, UMAP was unstable. PCA/*k*-means identified 3 clusters: A′ (n=38, 64.4%); B′ (4, 6.8%); and C′ (17, 28.8%). Only C′ demonstrated significantly worse OS (mOS 54.0 months; HR 4.05, p=0.030). "High risk-MCN" (C′) was enriched for high ISS and driven by % FLC reduction at 3 months, baseline eGFR, and baseline LDH.

DISCUSSION Unsupervised ML analysis of real-world NDMM data identified 2 MCN-enriched phenotypes: "control-like" with increased survival and "high risk" with inferior outcomes. Within the MCN subgroup, a third phenotype with high mortality was identified. Cluster-based groupings showed superior OS discrimination vs ISS, R-ISS, and Kaiser HRCA. These findings suggest MCN heterogeneity can be detected at diagnosis and further defined with early treatment data. To our knowledge, this is the first ML framework to detect latent MCN phenotypes and stratify their survival risk vs real-world controls. Future directions include model refinement with larger datasets, independent prospective validation, and development of model-based tools for clinical integration.

https://doi.org/10.1182/blood-2025-5756