The 65th ASH Annual Meeting Abstracts

POSTER ABSTRACTS

903.HEALTH SERVICES AND QUALITY IMPROVEMENT -MYELOID MALIGNANCIES

Machine Learning Approach to Understand Real-World Treatment in Patients with Higher-Risk Myelodysplastic Syndromes

Vandana Priya[1], Vivek P. Vaidya[1], Smita Agrawal[1], Neeraj Singh[1], Kaveri Chatra[1], Dhaval Parmar[1], Raymond Yan[1], Rahul K. Das[1], Mahnoush Haririfar[1], Peter McMahon[2], Mellissa Williamson[2], Islam Sadek[2], Cosmina Hogea[2]

[1] ConcertAI, Cambridge, MA
[2] Gilead Sciences, Inc., Foster City, CA

Background: Higher-risk myelodysplastic syndromes (HR-MDS) is a severe and aggressive form of MDS that represents a group of rare and heterogeneous hematologic disorders that primarily affect the bone marrow and blood cells. Treatment typically involves a combination of supportive care and active treatment with targeted therapies, hypomethylating agents, chemotherapy, and hematopoietic stem cell transplant. In this study, we have evaluated real-world treatment patterns in patients with HR-MDS with machine-learning (ML) models to identify clinical and demographic features that may impact decisions about which patients receive systemic therapies and when the therapies are implemented.

Methods: We identified 821 patients aged ≥18 years with HR-MDS diagnosed between January 2015 and April 2022, using ConcertAI's RWD360 database linked with open claims data. RWD360 consists of structured records from US-based oncology electronic health record systems. Demographic and clinical data up to 1 year prior to the date of HR-MDS diagnosis were used to develop a family of XGBoost-based ML models to investigate drivers of treatment. We created five models: a time-to-therapy model that predicts patient-level hazard ratios and four classification models that predict likelihood of therapy in the first 2, 3, 6, and 12 months, respectively, following HR-MDS diagnosis (therapy yes/no models). Model performance was evaluated using Harrell's concordance index (C-index) and Akaike information criterion (AIC) for the time-to-therapy model and area under curve (AUC) for the therapy yes/no models. Key features included laboratory values, Charlson comorbidity index (CCI), and prescribed medications. To better capture the temporal characteristics of a patient's disease state evolution, clinical features were organized into time windows alongside the variations within them. We used recursive feature elimination to bring the total features used in the model to 50 to enhance the interpretability and compactness. Five-fold cross-validation (data was split 5 ways; for each fold, a different 20% is left out for validation, while 80% of the data was used to train the model) on the full 821-patient cohort was performed to train and evaluate the model.
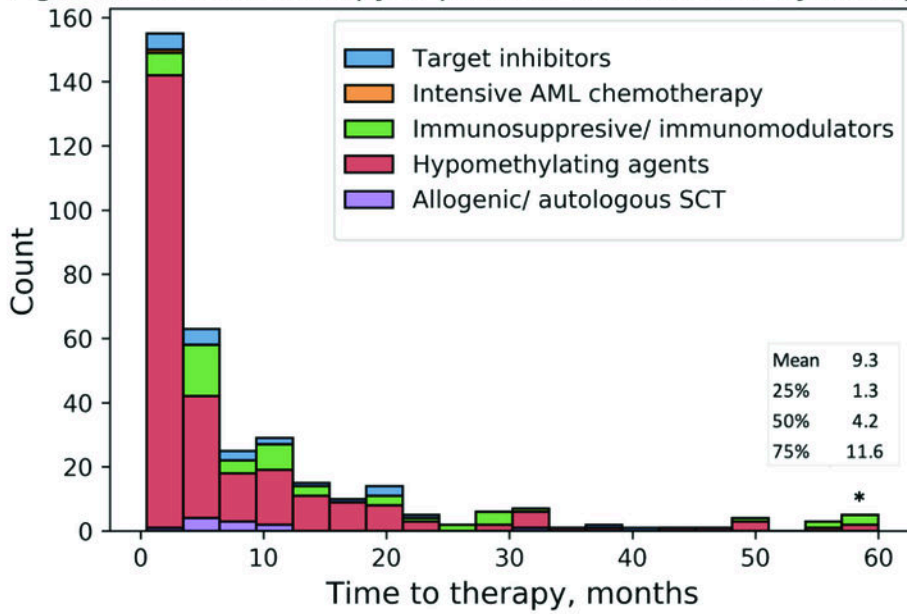
Results: The median age in this patient cohort was 73 years (range = 21-88 years); approximately 60% were male, and 66% patients were treated in community oncology practices. Only 42.5% patients were recorded to ever receive systemic active therapy, and of those, the median time to therapy was 4.2 months (Figure 1). Patients had a median hemoglobin concentration of 9.4 g/dL (range = 4.4-18.9 g/dL), median CCI of 1 (range = 0-10), and median polypharmacy of 10 (computed on patients who had at least 1 concomitant medication (49%); range = 1-93) around time of diagnosis (assessed 1 year prior, and up to 30 days after diagnosis date).

Model performance for the time-to-therapy model (C-index = 0.78; AIC = 3314.2) and for the therapy yes/no models (AUC = 0.7-0.74) were evaluated. A key predictive feature across models was the minimum laboratory value of white blood cell (WBC) count near the point of diagnosis; low WBC values around time of diagnosis was consistently a strong predictor for receiving systemic therapy and receiving it earlier. Two other top predictors of systemic therapy and its timing were the rate of fall in measurements of blood glucose and the rate of fall of hemoglobin during the 180 to 30 days before diagnosis. Table 1 details model performance and the top 3 features impacting time and likelihood of therapy for all 5 models.

Conclusion: We have built a family of data-driven ML models for patients with HR-MDS, based on data collected in routine clinical practice, towards predicting time-to and probability-of first-line systemic therapy with reasonable accuracy driven by patient characteristics. A majority of patients diagnosed with HR-MDS in this study appeared to receive delayed or no systemic therapy in their treatment journey. To bridge gaps in care and facilitate timely treatment in eligible patients, data-driven predictive modeling that take into account detailed patient-level factors is critical. Further analysis on outcomes of untreated patients linked with their key clinical characteristics identified using these models may provide opportunities to improve clinical care by implementing strategies for treatments based on patient-level factors.

Disclosures Priya: *ConcertAI:* Current Employment; *Gilead Sciences, Inc.:* Other: Support for the present manuscript. **Vaidya:** *General Electric:* Patents & Royalties; *ConcertAI:* Current Employment; *Gilead Sciences, Inc.:* Other: Support for the present manuscript. **Agrawal:** *ConcertAI:* Current Employment, Current equity holder in private company; *Gilead Sciences, Inc.:* Other: Collaborator and client supporting the study monetarily. **Singh:** *ConcertAI:* Current Employment; *Gilead Sciences, Inc.:* Other: Support for the present manuscript. **Chatra:** *ConcertAI:* Current Employment; *Gilead Sciences, Inc.:* Other: Support for the present manuscript. **Parmar:** *ConcertAI:* Current Employment; *Gilead Sciences, Inc.:* Other: Support for the present manuscript. **Yan:** *ConcertAI:* Current Employment. **Das:** *PredxBio:* Consultancy; *ConcertAI:* Current Employment, Other: Owns stock or stock options in private company; Support for attending meetings and/or travel; *Aitia:* Other: Owns stock or stock options in private company. **Haririfar:** *ConcertAI:* Current Employment; *Gilead Sciences, Inc.:* Other: Grant or contract agreement. **McMahon:** *Gilead Sciences, Inc.:* Current Employment. **Williamson:** *Gilead Sciences, Inc.:* Current Employment, Current equity holder in publicly-traded company; *Amgen Inc.:* Current equity holder in publicly-traded company; *F. Hoffmann-La Roche AG:* Current equity holder in publicly-traded company. **Sadek:** *Gilead Sciences, Inc.:* Current Employment. **Hogea:** *Gilead Sciences, Inc.:* Current Employment.

Figure 1. Time to therapy in patients with HR-MDS by therapy type

Figure 1. Time to therapy in patients with HR-MDS by therapy type

Legend:
- Target inhibitors
- Intensive AML chemotherapy
- Immunosuppresive/ immunomodulators
- Hypomethylating agents
- Allogenic/ autologous SCT

Mean	9.3
25%	1.3
50%	4.2
75%	11.6

Y-axis: Count. X-axis: Time to therapy, months

* - Time to therapy greater than 60 months were mapped to 60 month
Abbreviations : AML, acute myeloid leukemia; SCT, stem cell transplant

Table 1. Model performance and top features

Model Outcome	Performance	Top 3 Features and impact on model prediction
Time to Therapy	C-Index = 0.68 AIC = 3314.2	Time to therapy prediction shortens when: 1. Minimum value of WBC count within 30 days of DX is *Low* 2. Mean of ECOG values observed within 30 days of DX is *Low* 3. Rate of fall in **Hemoglobin** during the 180 to 30 days before Dx is *High*
Therapy received within 2 months of Dx vs not received within 2 months of Dx	AUC = 0.74	Likelihood of therapy rises when: 1. Standard deviation of **weight** within 30 days of Dx is *High* 2. Minimum **WBC** count within 30 days of Dx is *Low* 3. Rate of fall in measurements of blood **glucose** during the 180 to 30 days before Dx is *High*
Therapy received within 3 months of Dx vs not received within 3 months of Dx	AUC = 0.72	Likelihood of therapy rises when: 1. Rate of fall in measurements of blood **glucose** during the 180 to 30 days before Dx is *Low* 2. Standard deviation of weight within 30 days of Dx is *High* 3. Minimum **WBC** count within 30 days of Dx is *Low*
Therapy received within 6 months of Dx vs not received within 6 months of Dx	AUC = 0.7	Likelihood of therapy rises when: 1. Minimum **WBC** count within 30 days of Dx is *Low* 2. Rate of fall in measurements of blood **Glucose** during the 180 to 30 days before Dx is *High* 3. Mean of systolic blood pressure within 30 days of Dx is High
Therapy received within 12 months of Dx vs not received within 12 months of Dx	AUC = 0.72	Likelihood of therapy rises when: 1. Minimum **WBC** count within 30 days of Dx is *Low* 2. Mean **Platelet** count within 30 days of Dx is *Low* 3. Mean of ECOG values observed within 30 days of Dx is *Low*

AIC, Akaike information criterion; AUC, area under the curve; C-index, Harrell's concordance index; Dx, diagnosis; ECOG, Eastern Cooperative Oncology Group; WBC, white blood cell.

Figure 1

https://doi.org/10.1182/blood-2023-182241