#1

MDS: Use Case Overview & Key Results

Elisabetta Sauta, PhD

Humanitas Research Hospital – AI Center

elisabetta.sauta@humanitas.it

Overview of MDS Use Case

Challenges of Rare Hematological Diseases

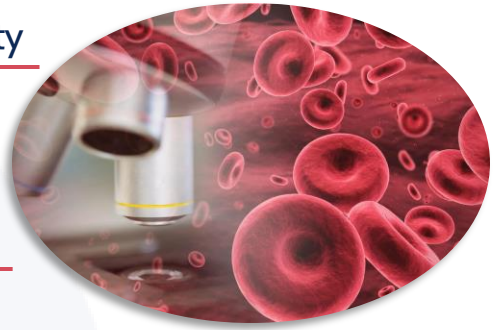
In MDS, defects in the bone marrow microenvironment and the hematopoietic cells lead to ineffective hematopoiesis. Its classification relies on morphological features and genetic abnormalities (defined by the Revised International Prognostic Scoring System).

FEATURES

Clinical Heterogeneity

Varied Genomic Background

Multifactorial Disease



MANAGEMENT ISSUES

Current prognostic classification systems fail to capture individual patient heterogeneity

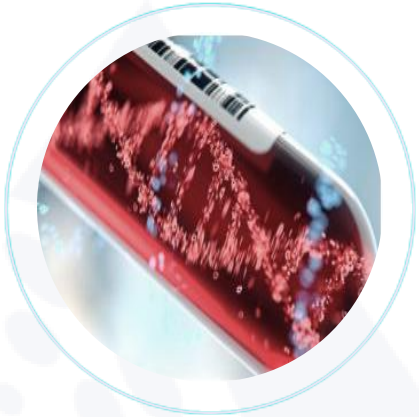
The majority of patients fail first-line therapies

Hematopoietic stem cell transplantation (HSCT) is the **only** potentially curative option but **not all patients** are eligible

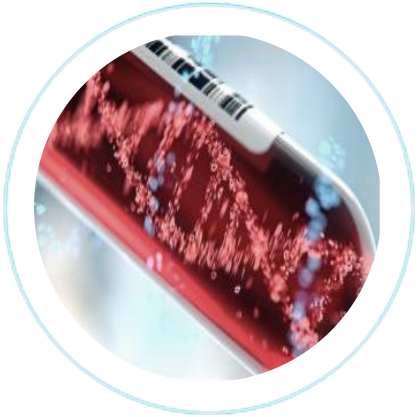
Personalized Risk-adapted Treatment Strategy

GenoMed4All Mission

Development of AI Solutions to Improve MDS Clinical Management through a Personalized Precision Medicine Approach



MDS Prevention Based on Genomic Screening



Omics-based Classification and Prognosis of MDS

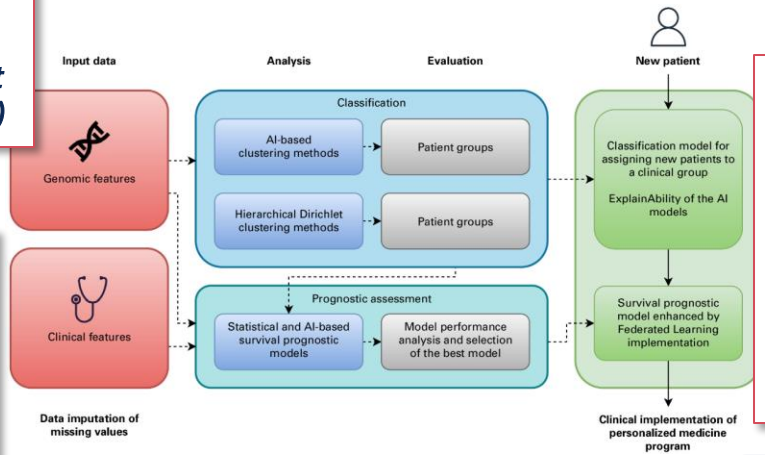
Omics-based Clinical Decision Making in MDS

MDS Prevention Based on Genomic Screening

MOSAIC: An Artificial Intelligence-based Framework for Multi-Modal Analysis, Classification and Personalized Prognostic Assessment in Rare Cancers

Data-driven, harmonized classification system for MDS: a consensus paper from the International Consortium for Myelodysplastic Syndromes, ***Lanino L et al. The Lancet Hematology. 2024 (in press)***

A Molecular-Based Ecosystem to ImprovePersonalized Medicine in Patients with ChronicMyelomonocytic Leukemia (CMML)
Oral presentation at the 66th American Society of Hematology (ASH) Conference (7-10 Dec. 2024)



D'Amico S et al. JCO Clin Cancer Inform. 2024. 8:e2400008. doi: 10.1200/CCI.24.00008.

Artificial-Intelligence, Data-Driven, Comprehensive Classification of Myeloid Neoplasms Based on Genomic, Morphological and Histological Features: **the TITAN Study**
Oral presentation at the 66th American Society of Hematology (ASH) Conference (7-10 Dec. 2024)

Omics-based Classification and Prognosis of MDS

Real-World Validation of the Molecular International Prognostic Scoring System (IPSS-M) for MDS Risk Stratification

The IWG for prognosis in MDS* proposed a new clinical-molecular prognostic model, the Molecular International Prognostic Scoring System [**IPSS-M**] to improve the prediction of clinical outcomes of the currently available tool (Revised International Prognostic Scoring System [IPSS-R]).

IPSS-R

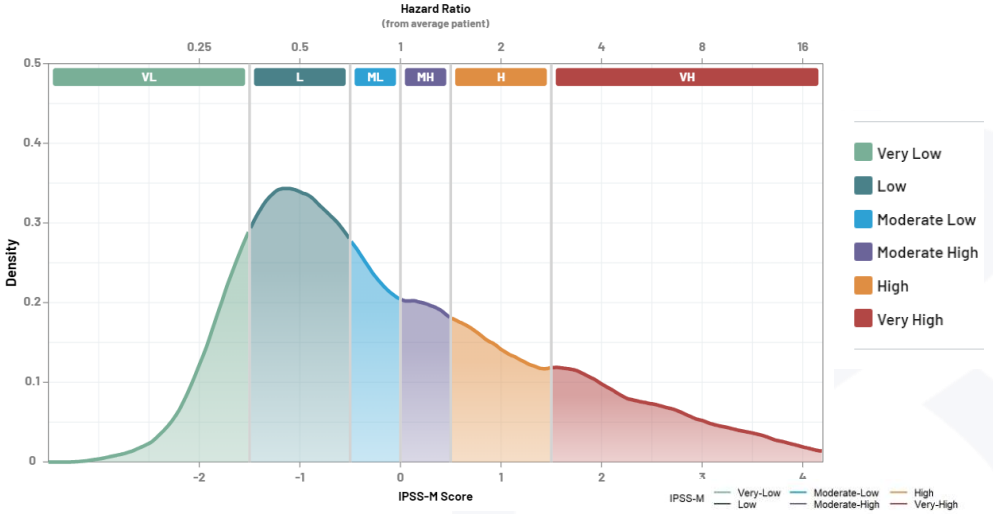
IPSS-M

Age, sex

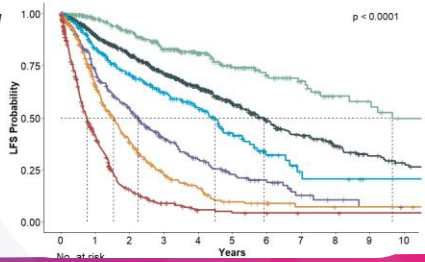
Blood Parameters

Cytogenetic Alterations

Somatic Mutations on 31 MDS-related genes



*Bernard E et al. 2022. NEMJ

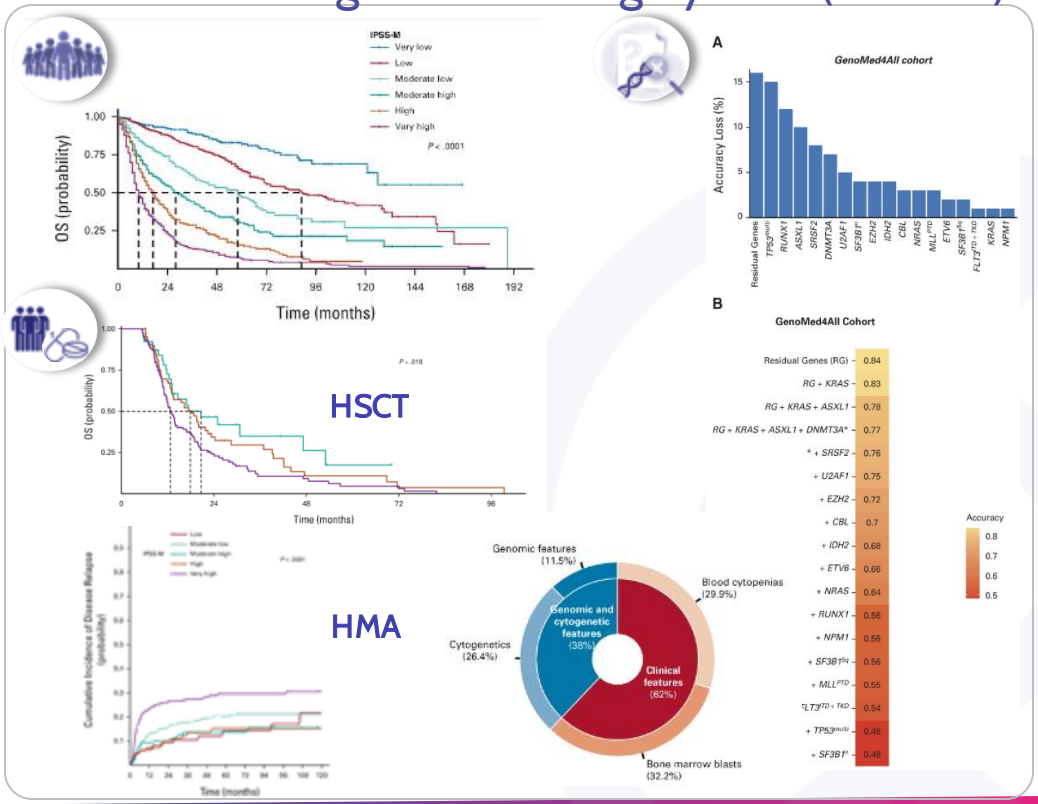


Omics-based Classification and Prognosis of MDS

Real-World Validation of the Molecular International Prognostic Scoring System (IPSS-M) for MDS Risk Stratification

On behalf of the GenoMed4All consortium*, we provided **extensive validation of the IPSS-M** in a real-world MDS cohort (n=2,876), addressing its clinical implementability by:

- Validating the IPSS-M prognostic value
- Investigating the IPSS-M prognostic power in patients receiving disease-modifying treatments
- Testing the prediction's accuracy with missing genomic information



*Sauta E et al. J Clin Oncol. 2023;41(15):2827-2842. doi: 10.1200/JCO.22.01784

Omics-based Classification and Prognosis of MDS

IPSS-M Clinical Implementability

Validating the IPSS-M prognostic value

Compared to IPSS-R, the IPSS-M resulted in **improved prognostic accuracy** across all clinical endpoints.

Investigating the IPSS-M prognostic power in patients receiving disease-modifying treatments

IPSS-M significantly **improved the risk prediction of relapse** and of the **probability of post-transplantation survival**, helping the identification of patients with high risk of transplantation failure.

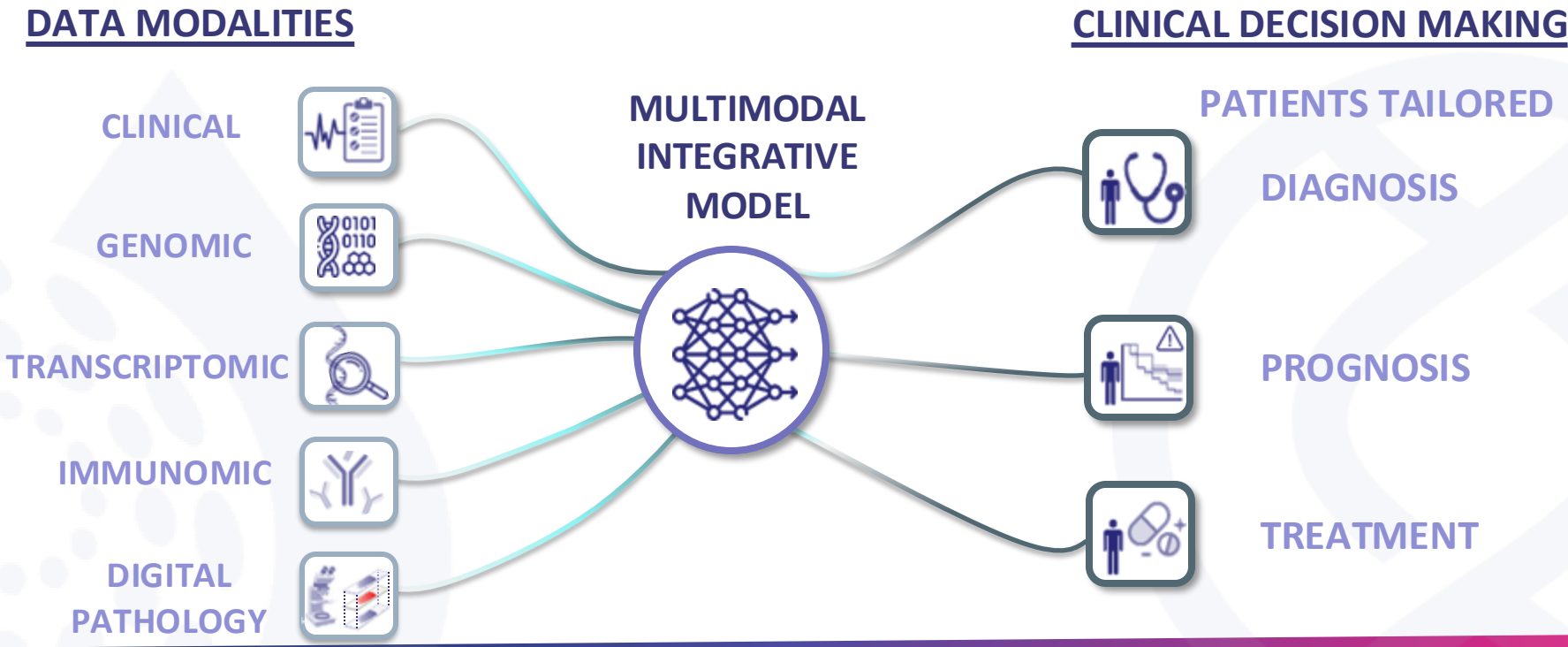
IPSS-M **failed to stratify individual probability of response**; additional factors other than gene mutations can be involved in determining sensitivity to HMA.

Testing the prediction's accuracy with missing genomic information

Testing the robustness of IPSS-M when molecular information was missed, we defined a **minimum set of 15 relevant genes** ensuring a **risk prediction accuracy greater than 70%**.

Omics-based Classification and Prognosis of MDS

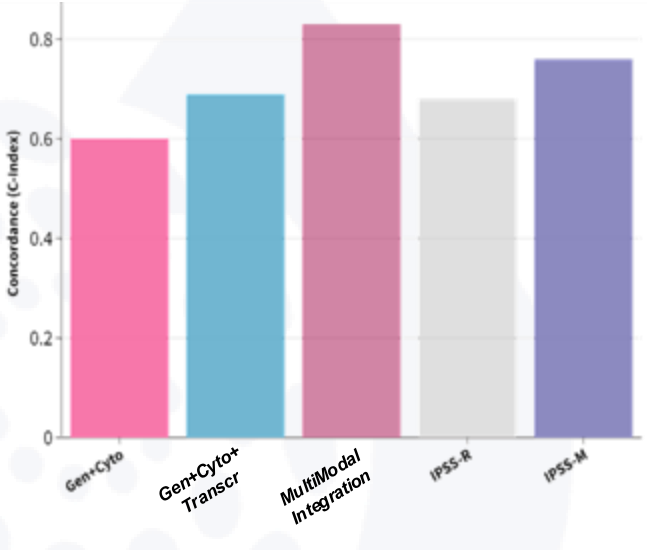
Multi-omics Analysis For Personalized Medicine in MDS

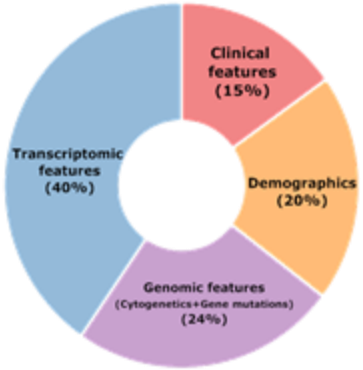


Omics-based Classification and Prognosis of MDS

Combining Gene Mutation with Transcriptomic Data Improves Outcome Prediction in MDS

We performed an integrative analysis using conventional statistical methods (Gerstung M et al. Nat Comm 2015) to evaluate the prognostic contribution of cytogenetic, transcriptomic, genomic, clinical and demographic features in predicting clinical outcomes in MDS. We assessed the relative contribution of each data layer comparing the obtained accuracy with the current standard prognostic scoring systems, IPSS-R and IPSS-M.





Pros: Accessible interpretation

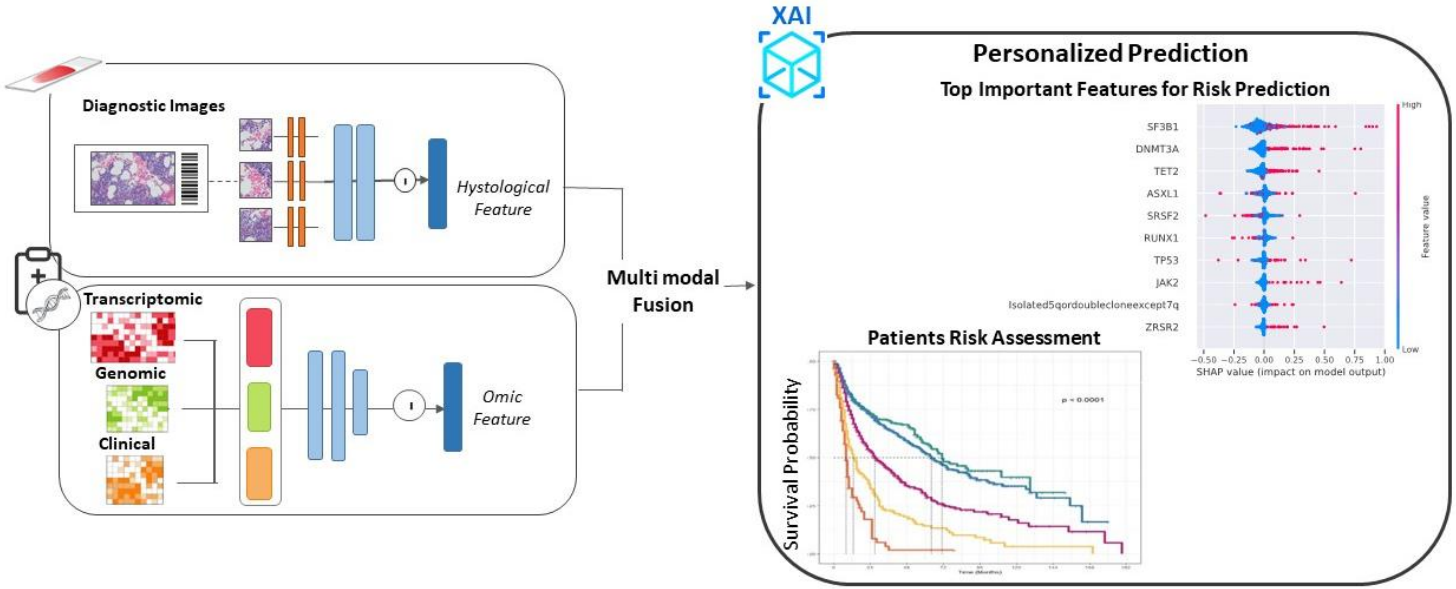
Cons:

- A limited number of input features
- «Manual» features extraction
- Non-scalable model

Sauta E et al. Blood 2023; 142 (Supplement 1): 1863. doi: https://doi.org/10.1182/blood-2023-186222

Omics-based Classification and Prognosis of MDS

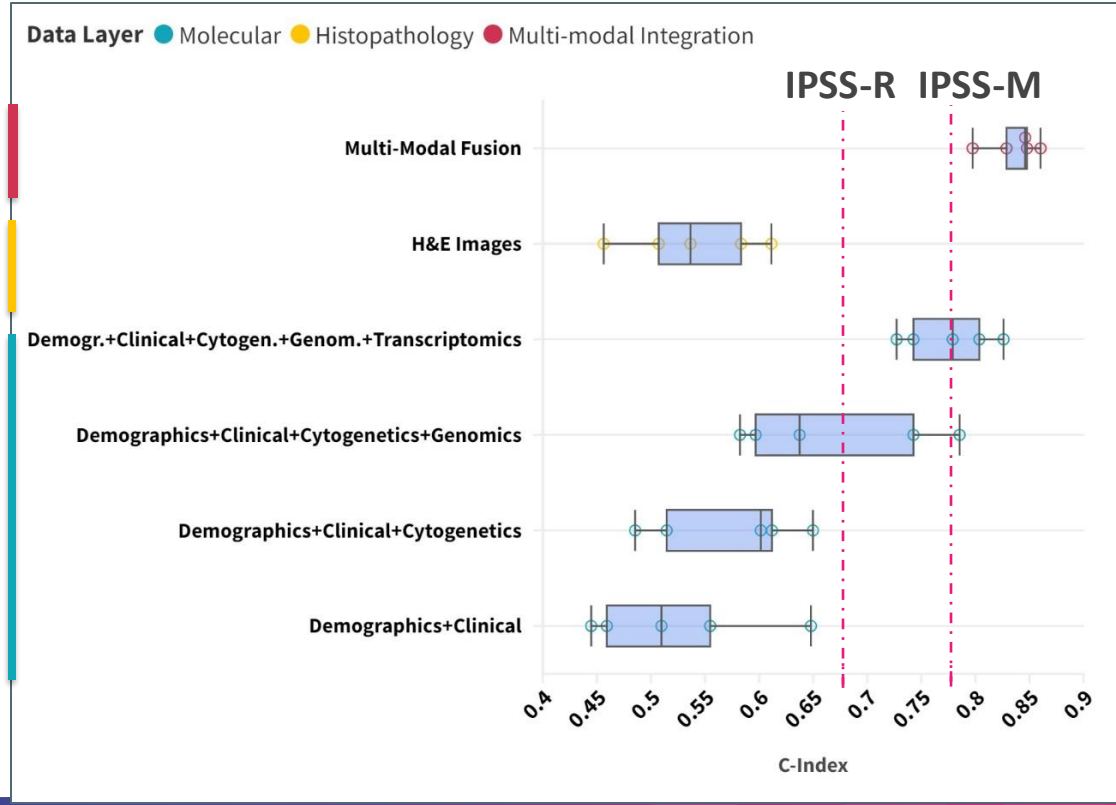
Multimodal AI-driven Platform for Precisione Medicine in MDS



Omics-based Classification and Prognosis of MDS

Multimodal AI-driven Platform for Precisione Medicine in MDS

MDS Cohort

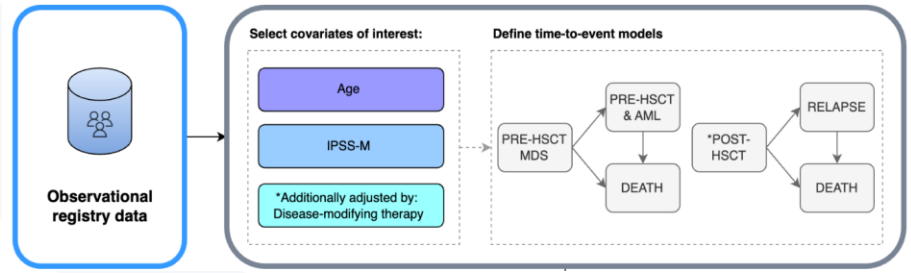


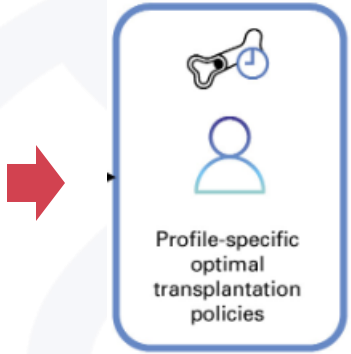
Sauta E et al. accepted as oral presentation at the 66th ASH conference (7-10 Dec.2024)

Omics-based Clinical Decision Making in MDS

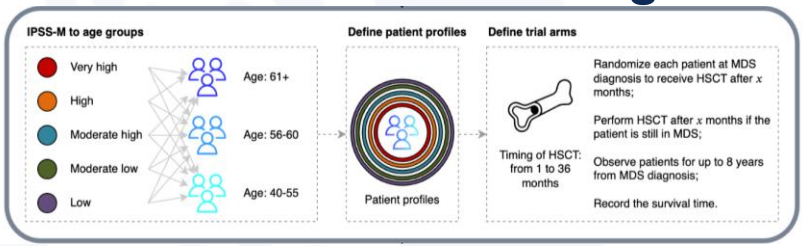
Clinical and Genomic-Based Decision Support System to Define the Optimal Timing of Allogeneic Hematopoietic Stem-Cell Transplantation in MDS

STEP 1 – Model of the disease natural history

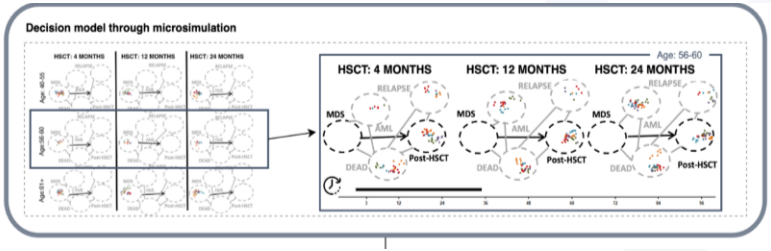




STEP 2 Simulation of the target trial

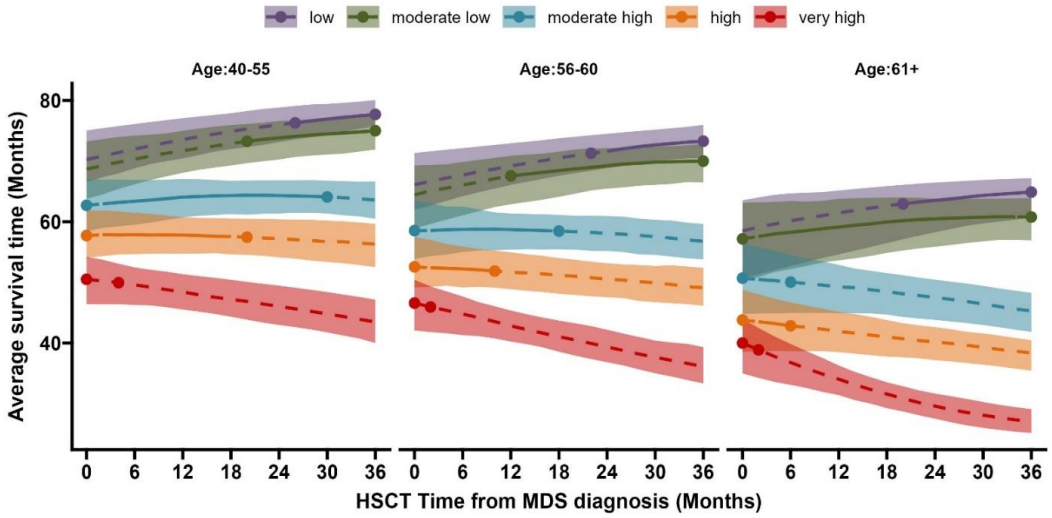


STEP 3 Microsimulation scenario



Omics-based Clinical Decision Making in MDS

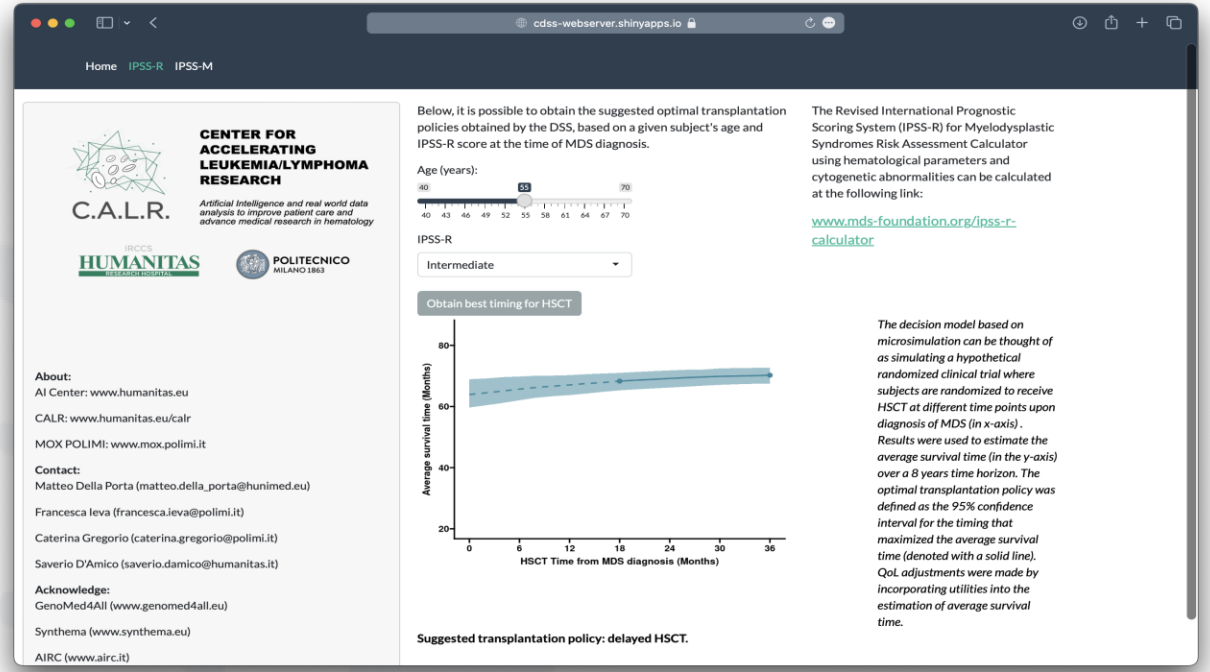
Clinical and Genomic-Based Decision Support System to Define the Optimal Timing of Allogeneic Hematopoietic Stem-Cell Transplantation in MDS



Under an IPSS-M based policy, in the patients with either **low- and moderate-low risk** benefited from a **delayed transplantation** policy, while in those belonging to **moderate-high, high- and very-high risk** categories **immediate transplantation** is recommended

Omics-based Clinical Decision Making in MDS

Clinical Decision Support System for Transplantation in MDS - WEB TOOL



Gregorio C et al. JCO Clin Cancer Inform. 2024;8:e2300205. doi: 10.1200/CCI.23.00205

Tentori C et al. J Clin Oncol. 2024. 42(24):2873-2886. doi: 10.1200/JCO.23.02175

Acknowledgements

European Reference Network
for rare or low prevalence complex diseases
Network Hematological Diseases (ERN EuroBloodNet)

Co-funded by the European Union

This project is supported by the European Reference Network on Rare Haematological Diseases (ERN-EuroBloodNet)-Project ID No 101085717. ERN-EuroBloodNet is partly co-funded by the European Union within the framework of the Fourth EU Health Programme.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Health and Digital Executive Agency (HaDEA). Neither the European Union nor the granting authority can be held responsible for them.

GENOMED4ALL

Funded by the European Union

GenoMed4All has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017549.

Acknowledgements

SYNTHEMA

GENOMED4ALL

European Union Funding for Research & Innovation

IRCCS HUMANITAS RESEARCH HOSPITAL

HUMANITAS AI CENTER

CENTER FOR ACCELERATING LEUKEMIA/LYMPHOMA RESEARCH

C.A.L.R.

Consiglio Nazionale delle Ricerche

train

Thanks!

Any questions?

GenoMed4All & ERN-EuroBloodNet

Educational Program on AI in Hematology for an expert audience

Follow us!

genomed4all.eu

@genomed4all

/genomed4all

eurobloodnet.eu

@ERNEuroBloodNet

/ERNEuroBloodNet

GenoMed4All & ERN-EuroBloodNet

Educational Program on AI in Hematology for an expert audience

#1

MDS: Use Case Overview & Key Results

Saverio D'Amico

Humanitas Research Hospital – AI Center

saverio.damico@humanitas.it

Synthetic Data & Digital Twin

Overview and applications in MDS

Increased access to real-world evidence (RWE) data is needed to accelerate innovation in hematology

97% of healthcare data remains unused

Main reasons

- Privacy limitations (GDPR)
- Lack of data harmonization from different sources
- Data are not structured and are dispersed

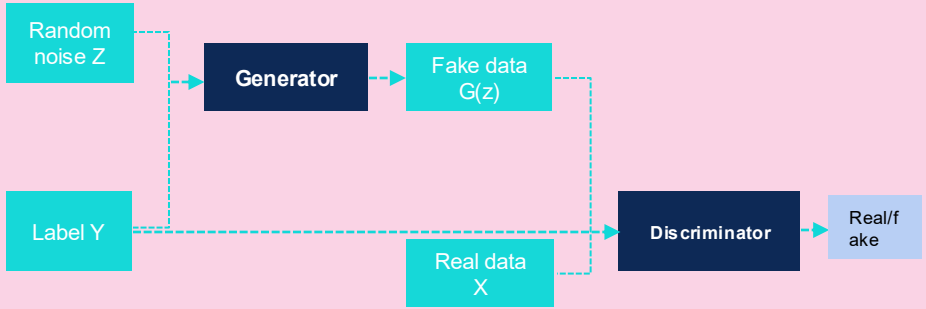
Source: Deloitte 2023

Synthetic data to capture RWE in hematology

Synthetic data are artificial data generated by an algorithm trained to learn all the essential characteristics of a real dataset:

- The new data are neither a copy nor a representation of the real data
- Since they are not real data, they are not regulated by particular limitations; therefore, they can be easily accessed and shared

Conditional Generative Adversarial Networks architecture

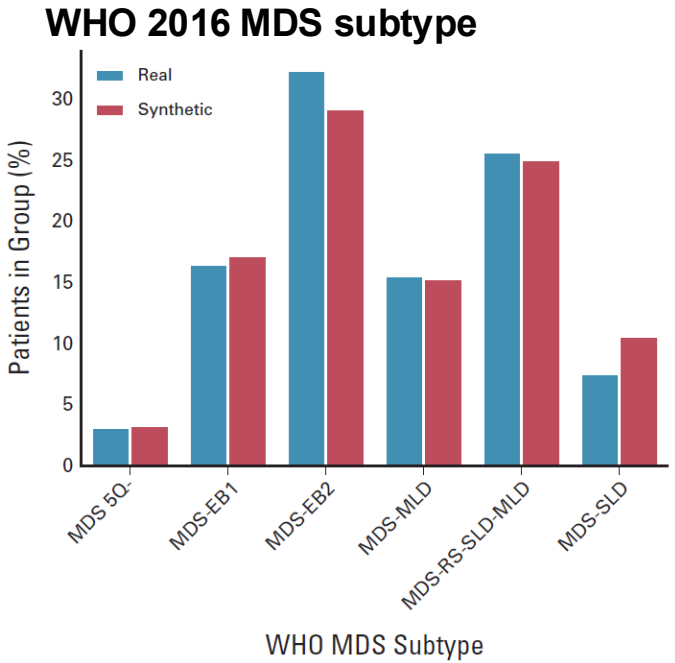


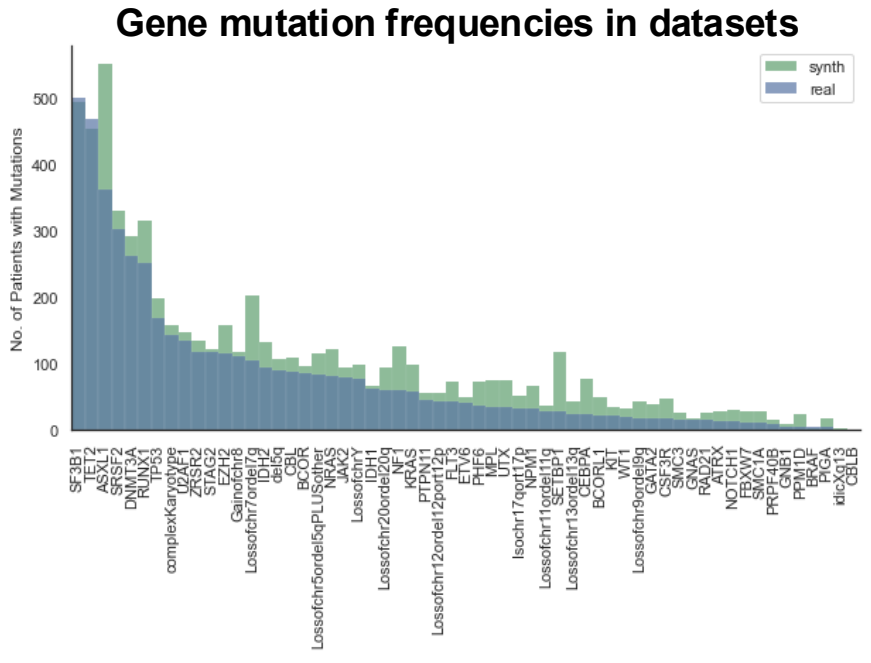
Properties and possible applications

- Data sharing (privacy/GDPR)
- Classes balance and resolution of missing information (data harmonization)
- Data augmentation
- Algorithms training and validation
- Generation of new evidence

Source: D'Amico et al, JCO CCI, 2023

Synthetic vs. real patients: Comparison of clinical and molecular features in MDS



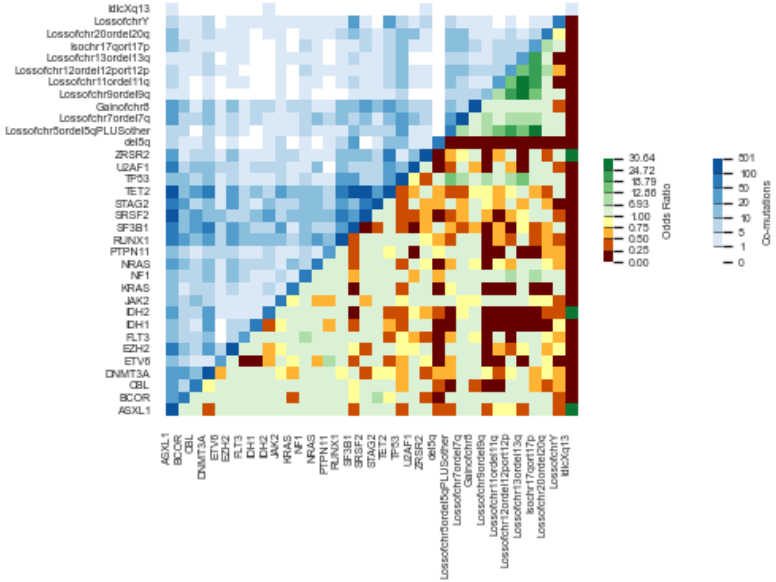


Source: D'Amico et al, JCO CCI, 2023

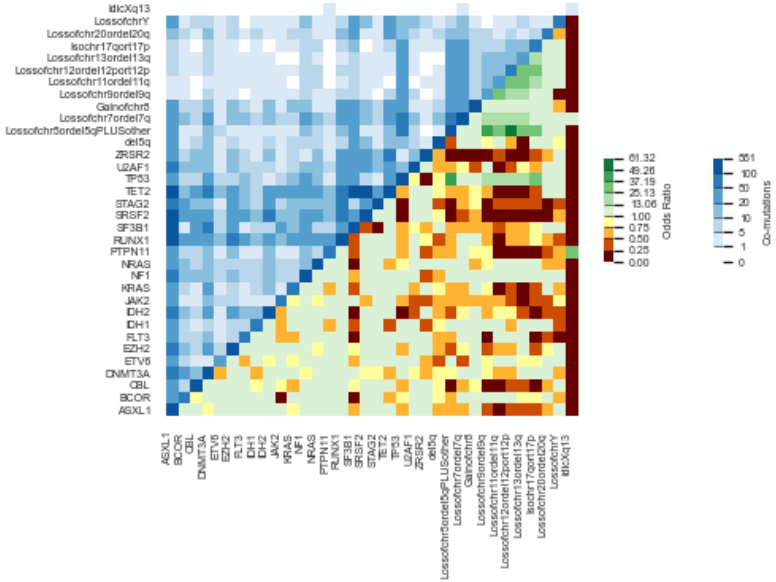
Synthetic vs. real patients: Comparison of clinical and molecular features in MDS

Pairwise associations among genes and cytogenetic abnormalities

Real data



Synthetic data



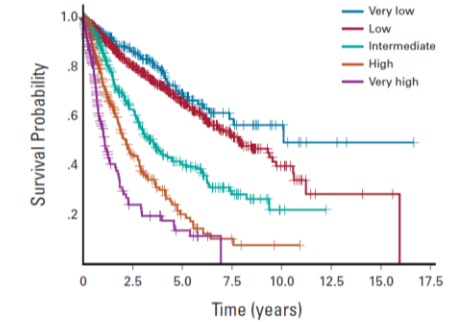
Source: D'Amico et al, JCO CCI, 2023

Synthetic vs. real patients: Comparison of clinical and molecular features in MDS

Probability of survival stratified by IPSS-R[1]

Data augmentation: From 2,043-7,118 patients[1,2]

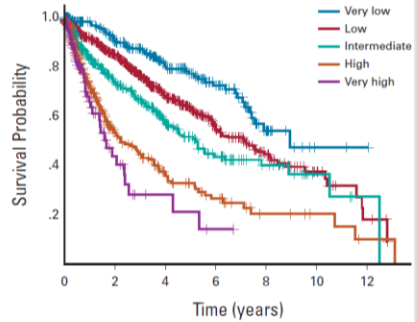
Real patients



No. at risk:

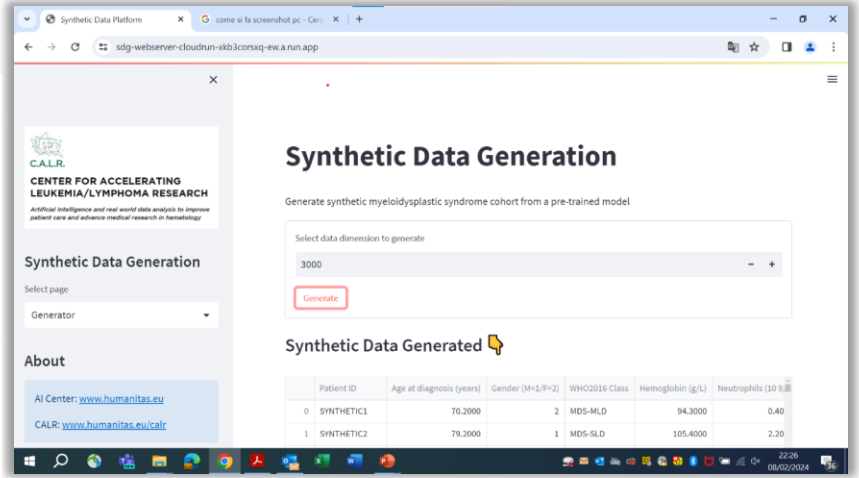
Very low	197	118	57	24	9	3	1	0
Low	578	329	174	68	18	3	2	0
Intermediate	328	126	56	21	2	0	0	0
High	281	55	14	4	1	0	0	0
Very high	215	16	6	0	0	0	0	0

Synthetic patients



No. at risk:

Very low	255	163	96	55	16	3	2
Low	655	379	194	90	42	16	4
Intermediate	451	170	87	38	17	4	1
High	406	82	39	19	7	4	1
Very high	134	13	4	1	0	0	0



Source: D'Amico et al, JCO CCI, 2023

Performance of synthetic data

2021 WHO guidance on ethics and governance of AI for health

We need to address three important topics for the right deployment of AI in hematology:[1]

- **Transparency of models:** interpretability and explainability
- **Reliability of models:** independent validation of generated AI models
- **Protection of data and data sharing:** Compliance with GDPR (EU)

Demographic, clinical, and survival data[2,3]	Genomic data[2,3]	All data[2,3]
		
92.1 %	**90.2 %**	**70.6 %**
SYNTHETIC CLINICAL FITNESS	SYNTHETIC GENOMIC FITNESS	PRIVACY PRESERVABILITY
Evaluated with distribution plot, Principal Component Analysis, and correlation matrices	Evaluated with mutation frequencies and pairwise association	Evaluated considering the possibility of tracing real data from synthetic ones

Source: D'Amico et al, JCO CCI, 2023

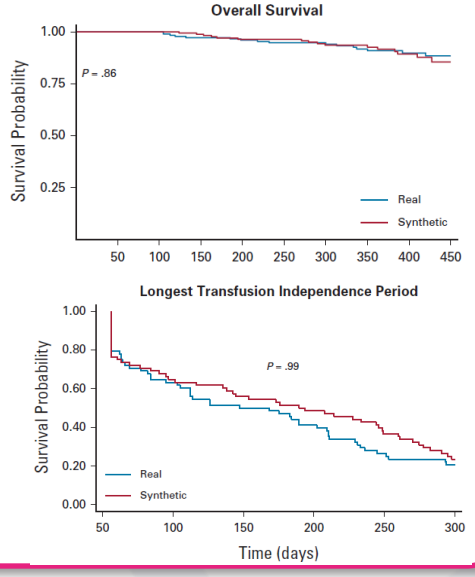
Synthetic Data (from RWE) to accelerate clinical research in hematology

Comparing endpoints of clinical trials using real and synthetic control arms

Real-world efficacy and safety of luspatercept in patients with transfusion-dependent anemia due LR-MDS-RS, who had an unsatisfactory response to or are ineligible for erythropoietin-based therapy: A multicenter study by FiSiM[1,2]

Clinical endpoint[1]	Real data	Synthetic data	p-value
RBC-TI ≥ 8 weeks 1-24	56 (31.5)	56 (31.5)	1.0
Longest transfusion independence period, weeks, median (range)	195 (56-490)	280 (56-490)	< 0.05
RBC-TI ≥ 8 weeks 1-48	68 (38.2)	61 (34.3)	0.50
RBC-TI ≥ 12 weeks 1-24	36 (20.2)	41 (23.0)	0.60
RBC-TI ≥ 12 weeks 1-48	51 (28.7)	46 (25.8)	0.63
Reduction ≥ 4 RBCs	62 (34.8)	63 (35.4)	1.0
Reduction ≥ 50%	77 (43.3)	72 (40.4)	0.66

Primary endpoints[1]



Source: D'Amico et al, JCO CCI, 2023

Limitations and pitfalls of randomized clinical trials (RCTs)

RCTs may have major limitations in some clinical scenarios, including:

Rare and ultrarare diseases (few patients available to be enrolled into a clinical trial)[1]

Patients with major unmet clinical needs in which best available therapy is not effective (ethical concerns to treat these patients with best available therapy)[2]

Patients with major unmet clinical needs in which most of the trials are failed (urgent need to accelerate the evaluation of the effectiveness of new treatment options)[1]

Diseases arising in elderly people (limited opportunity to be enrolled into a conventional clinical trial)[1]

Diseases in which the landscape of standard treatment is rapidly changing and therefore it is not easy to define an appropriate control arm (challenge in defining the study design)[2]

Synthetic data (from RWE) in clinical trials: Status of regulation and next plans

Alectinib obtained conditional EU approval as a treatment for lung cancer in **2017**, with acceptance of a synthetic control arm of 67 patients as a trial, thus accelerating drug's availability in the EU by 18 months[1]

Avelumab for the treatment of Merkel cell carcinoma, **approved in 2018**, used data from electronic medical records in a synthetic control arm[1]

Accelerated approval was obtained for blinatumumab for the treatment of leukemia from the **FDA in 2014 and the EMA in 2015**, using a comparator arm of historical data from 694 patients, based on 2,000 patient records for the Phase 2 study[1]

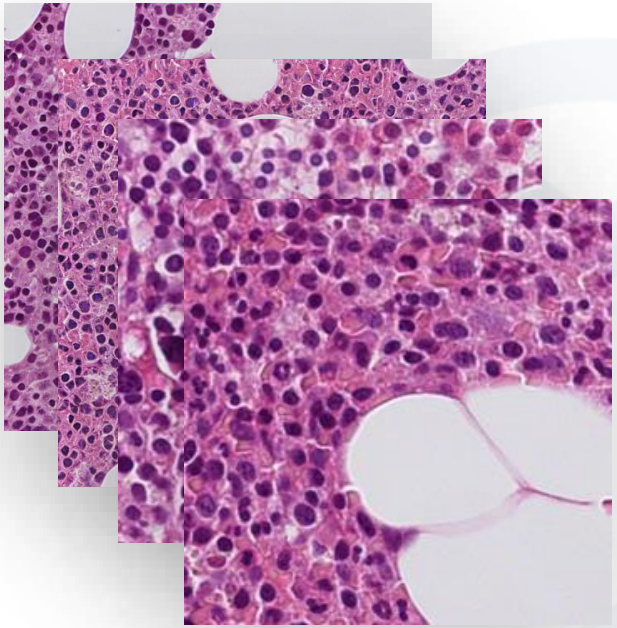
The FDA and the scientific community are forming an alliance for the **no-placebo initiative** to use external comparison arms to study new therapies for GI stromal tumor and other rare cancers, facilitating drug trials and regulatory approvals[2]

Synthetic data and images generation

	Patient ID	Age at diagnosis (years)	Gender (M=1/F=2)	WHO2016 Class	Hemoglobin (g/L)	Neutrophils (10 9/L)	Pl
0	SYNTHETIC1	68	2	MDS-MLD	135.6	0.6	
1	SYNTHETIC2	66.1	1	MDS-MLD	93.5	2.1	
2	SYNTHETIC3	37.8	2	MDS-SLD	96	1.9	
3	SYNTHETIC4	70.4	1	MDS-SLD	75	2.3	
4	SYNTHETIC5	77.2	2	MDS-SLD	84.8	4.7	
5	SYNTHETIC6	64.8	1	MDS 5Q-	118.2	1.3	
6	SYNTHETIC7	70.1	2	MDS-SLD	80.3	2.4	
7	SYNTHETIC8	71	1	MDS-MLD	87.3	0.7	
8	SYNTHETIC9	64.5	1	MDS 5Q-	89.5	7.4	
9	SYNTHETIC10	74.6	1	MDS-MLD	111.9	0.6	

	Patient ID	Age at diagnosis (years)	Gender (M=1/F=2)	WHO2016 Class	Hemoglobin (g/L)	Neutrophils (10 9/L)	Pl
0	SYNTHETIC1	68	2	MDS-MLD	135.6	0.6	
1	SYNTHETIC2	66.1	1	MDS-MLD	93.5	2.1	
2	SYNTHETIC3	37.8	2	MDS-SLD	96	1.9	
3	SYNTHETIC4	70.4	1	MDS-SLD	75	2.3	
4	SYNTHETIC5	77.2	2	MDS-SLD	84.8	4.7	
5	SYNTHETIC6	64.8	1	MDS 5Q-	118.2	1.3	
6	SYNTHETIC7	70.1	2	MDS-SLD	80.3	2.4	
7	SYNTHETIC8	71	1	MDS-MLD	87.3	0.7	
8	SYNTHETIC9	64.5	1	MDS 5Q-	89.5	7.4	
9	SYNTHETIC10	74.6	1	MDS-MLD	111.9	0.6	

Synthetic data generation



The importance of training

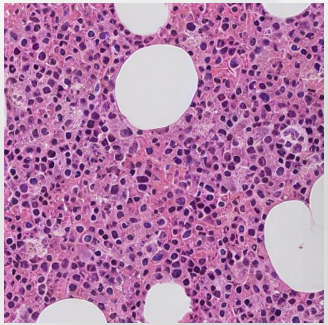
Prompt: Generate *cell images* of a patient of 80 years old with *acute leukemia*

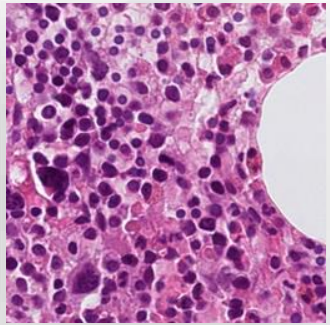
stability.ai

Generalistic Stable Diffusion Models

Prompt: Generate *cell images* of a patient of 80 years old with *acute leukemia*

Custom model trained on medical data



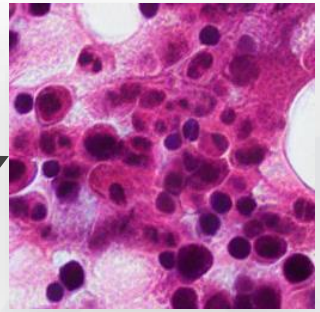


Generation of synthetic images from textual information

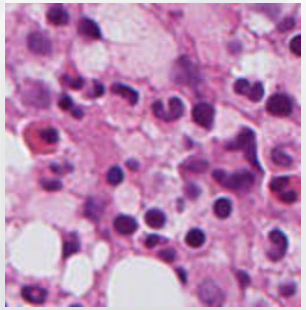
Clinical Text Description:

"BM cellularity 55% with preservation of the L:E ratio and by maturational progression of hematopoietic lines in the absence of relevant aspects of cytoarchitectural dysplasia. Proportion of CD34+ precursors <1%. Expression of p53 in less than 2% of the total cells. Absence of BM fibrosis."

Generative Model

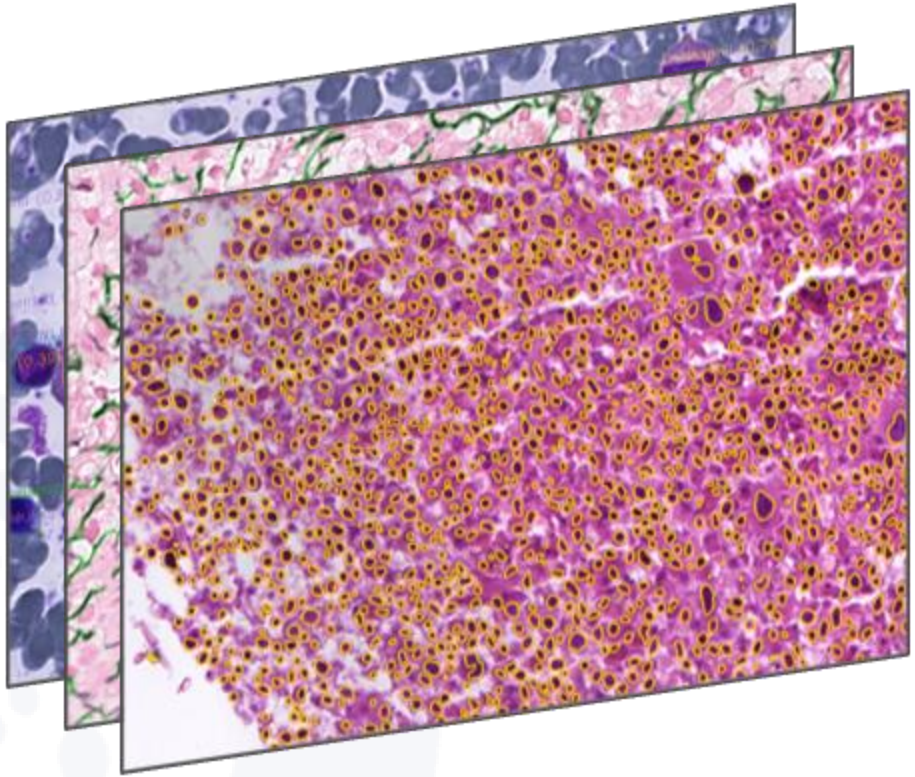


SYNTH IMAGE



REAL IMAGE

Clinical validation on imaging data: features extraction



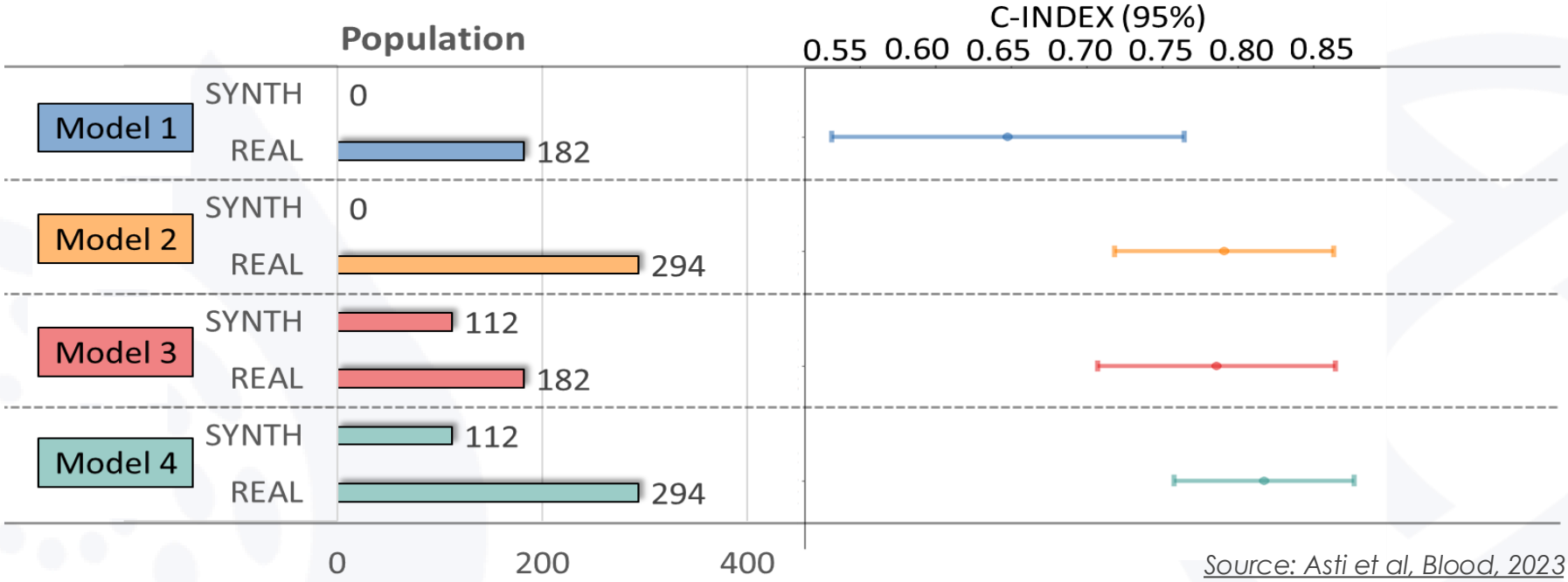
Morphological
Area
Perimeter
...

Color
Mean RGB
Mean HSV

Haralick
Entropy
Contrast

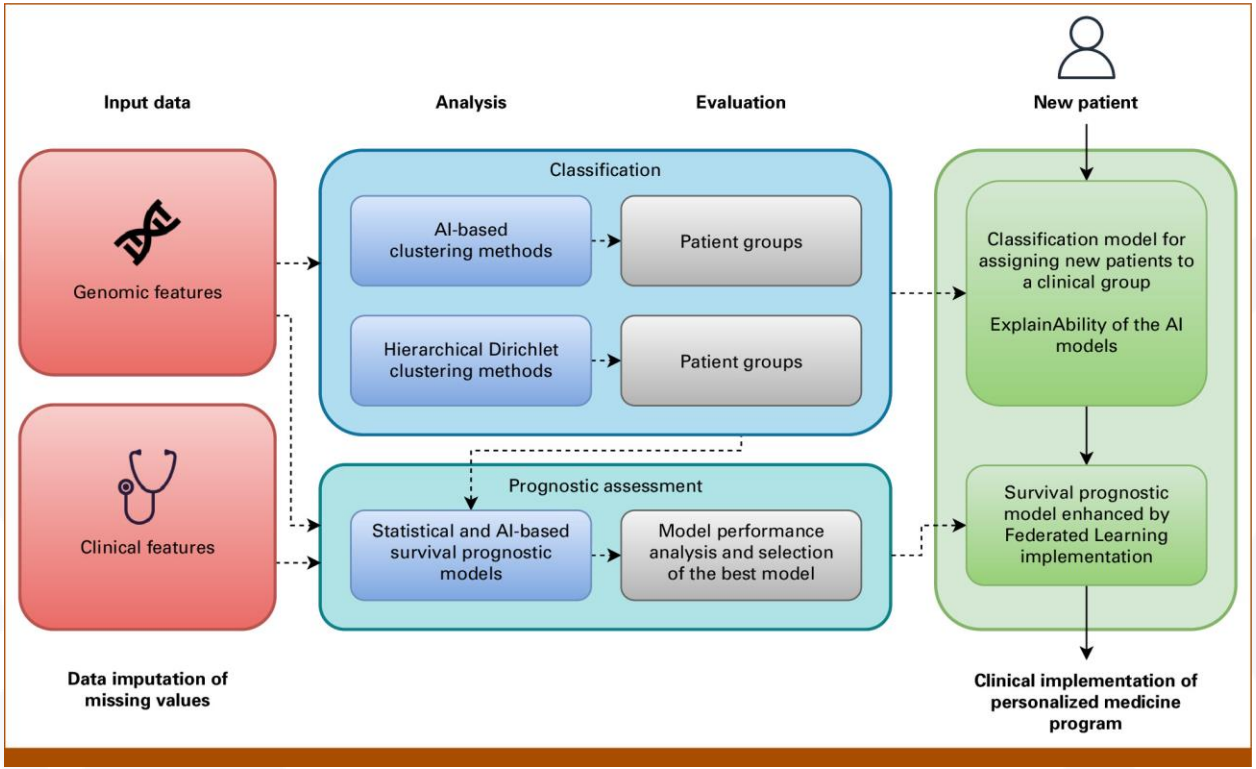
Clinical validation on imaging data: prognosis

Cox's proportional hazards model to predict individual probability of overall survival in patients affected with myeloid neoplasms.



Source: Asti et al, Blood, 2023

MOSAIC Framework

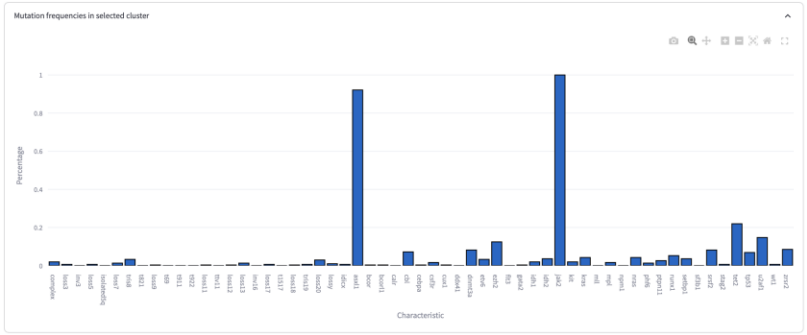


Source: D'Amico et al, JCO CCI, 2024

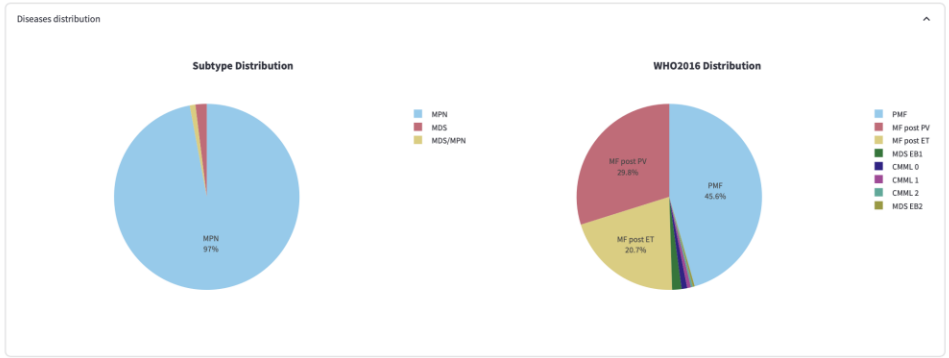
MOSAIC Framework for Comprehensive Classification of Myeloid Neoplasms Based on Genomic, Morphological and Histological Features: the «TITAN» study

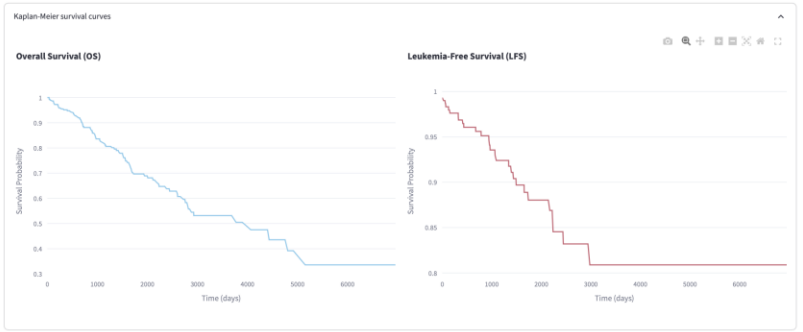
In the TITAN study, we retrospectively studied **20,054 patients** with MN in which clinical and morphological data, together with cytogenetics, mutational screening and outcome were available.

We included 7104 AML, 8410 MDS, 2986 MDS/MPN and 1554 MF. We used MOSAIC, an AI-based framework to define unsupervised clusters according to homogeneous morphological and genomic features.



Cluster n. 1





Abstract for 2024 ASH congress

The Lancet Haematology
Available online 9 October 2024
In Press, Corrected Proof What's this?

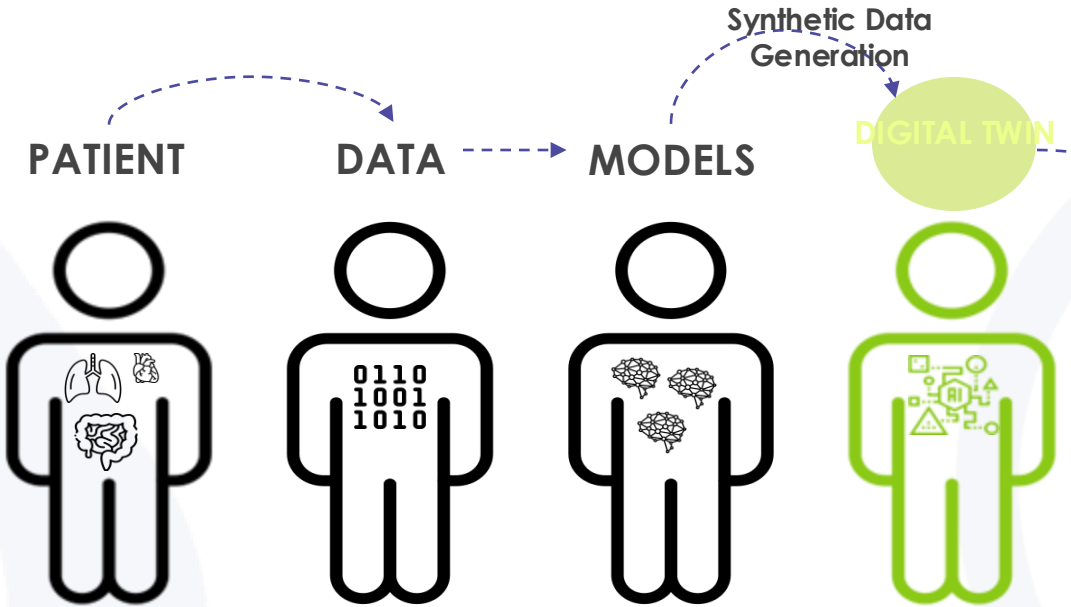
Review

Data-driven, harmonised classification system for myelodysplastic syndromes: a consensus paper from the International Consortium for Myelodysplastic Syndromes

Prof Rami S Komrokji MBBS [a] *, Luca Lanino MD [b c] *, Somedeb Ball MD [d] *, Jan P Bewersdorf MD [e] *, Monia Marchetti MD [f], Giulia Maggioni MD [b c], Erica Travaglino BSc [g], Najla H Al Ali MSc [a], Prof Pierre Fenaux MD [h], Prof Uwe Platzbecker MD [i], Valeria Santini MD [j], Maria Diez-Campelo MD [k], Avani Singh MD [l], Akriti G Jain MD [l], Luis E Aguirre MD [m], Sarah M Tinsley-Vance PhD [a], Zaker I Schwabkey MD [a], Onyee Chan MD [a], Zhouer Xie MD [a], Andrew M Brunner MD [m]...Prof Matteo G Della Porta MD [b c †]

From data collection to digital twin with synthetic data

Different data layers:
- Clinical
- Genomic
- Images



Prediction and simulation:
- Diagnosis
- Prognosis
- Therapy optimization

Digital Twin Platform for MDS

gemini-xkb3corsxq-ew.a.run.app

GEMINI: A Digital Twin Platform for Hematology (DEMO)

Select a page

Digital Twin

Select Patient Profile

Patient Profile

Patient profile: MDS SF3B1

Age

65

Sex

Male

Clinical Information

Hemoglobin

8,20

Neutrophils

1400,00

Platelets

257,00

Digital Twin Platfor

Select Clinical Question

Clinical Question

Diagnosis

Diagnosis

Prognosis

Treatment Strategy

gemini-xkb3corsxq-ew.a.run.app

GEMINI: A Digital Twin Platform for Hematology (DEMO)

Select a page

Digital Twin

Select Patient Profile

Patient Profile

Patient profile: MDS SF3B1

Age

65

Sex

Male

Clinical Information

Hemoglobin

8,20

Neutrophils

1400,00

Platelets

257,00

Digital Twin Platform for MDS patients

Select Clinical Question

Clinical Question

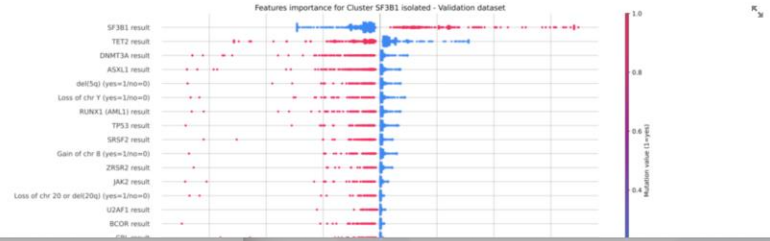
Diagnosis

Execute Model

According to the MDS classification model the patient is assign to the cluster:

SF3B1 isolated

Below are the features that drove the classification according to SHapley Additive exPlanations (SHAP)



ERN | EuroBloodNet

GENOMED 4ALL

Thanks!

Saverio D'Amico

Humanitas Research Hospital – AI Center

saverio.damico@humanitas.it