Immunotherapy in Hematological Malignancies 2023

May 18-19-20, 2023

Artificial Intelligence in Hematology

MATTEO DELLA PORTA, MD

Humanitas Research Hospital, Milan Italy

C.A.L.R.

HUMANITAS UNIVERSITY

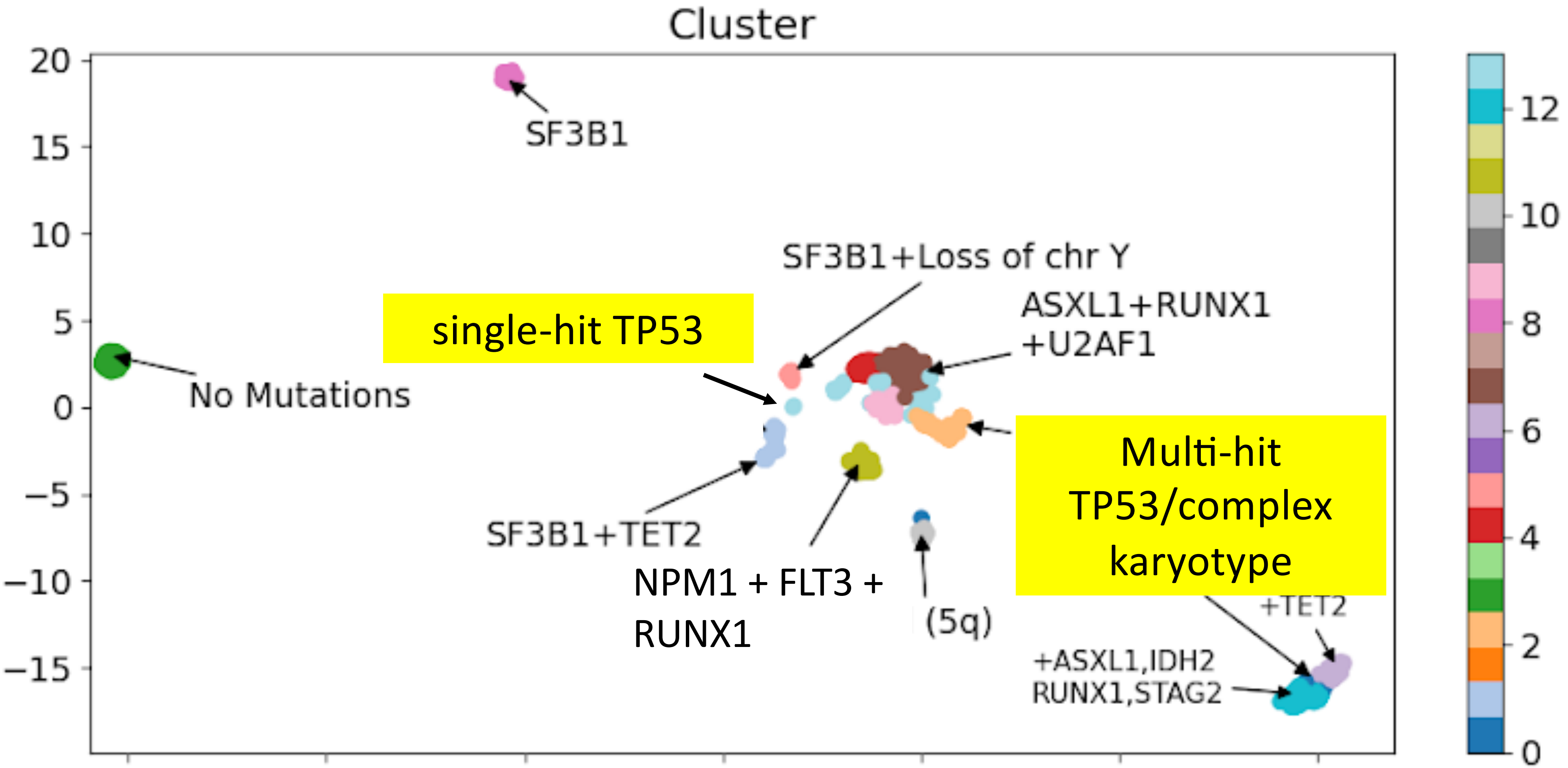
Artificial Intelligenece (AI) for precision medicine

1- Machine Learning

2- Generative AI

HUMANITAS

Molecular classification of MDS by machine learning methods (AI)



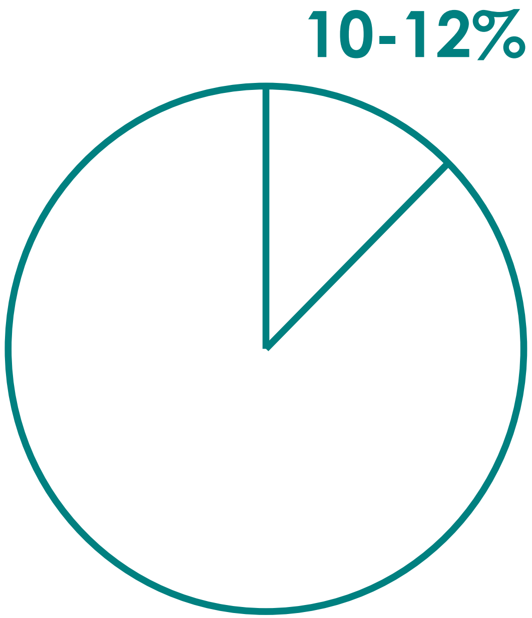
Bersanelli et al. J Clin Oncol 2021; D'Amico et al, J Clin Oncol CCI 2023, in press

GEN MED4ALL

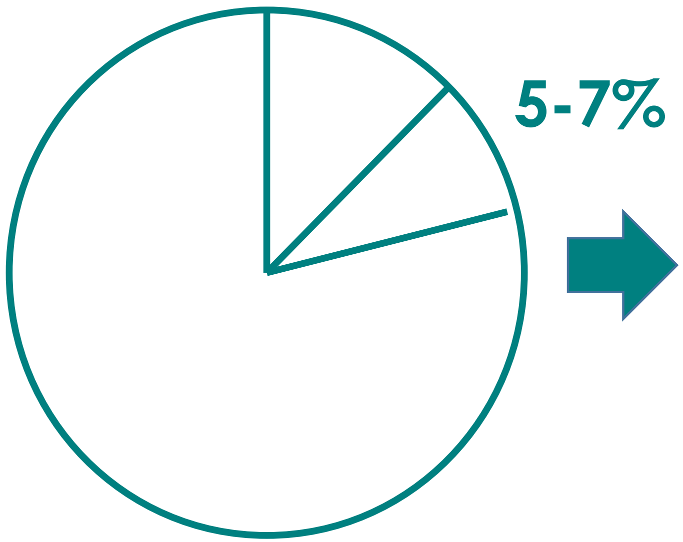
A new category of high-risk MDS is defined according to p53 dysfunction

SHAP (Shapley_Additive_Explanations) was used to explain the classification model by computing the contribution of each feature

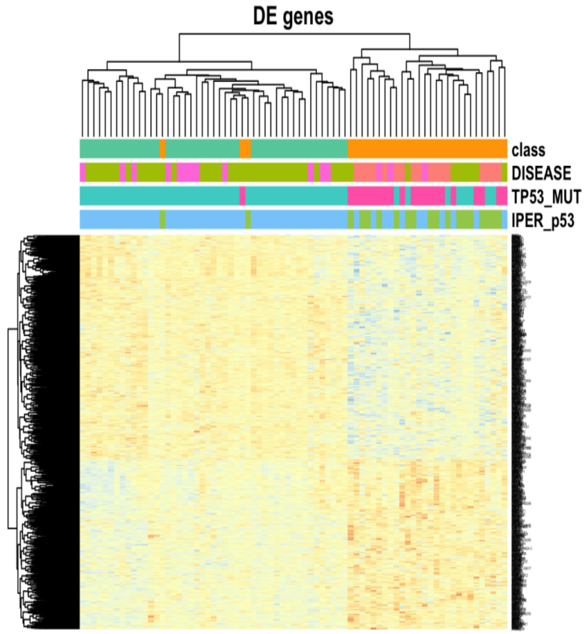
MDS with TP53-mutations



MDS with p53 dysfunction without TP53 mutations



RNA sequencing

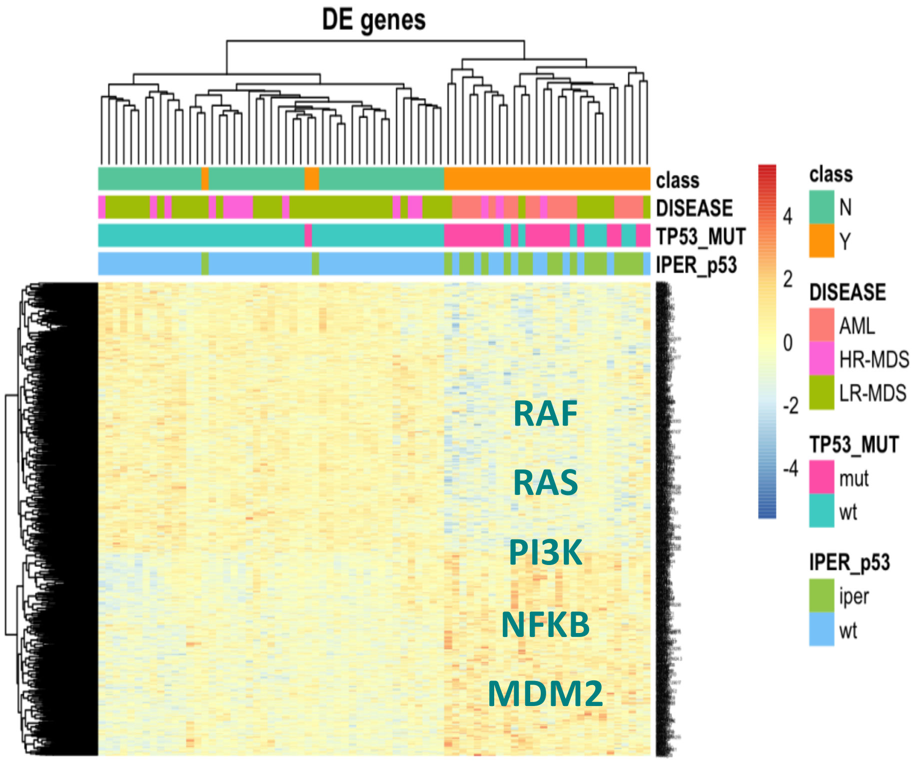


Bersanelli et al. J Clin Oncol 2021
Riva et al. Blood 2022;140:4001–4; Zampini et al, manuscript in preparation

HUMANITAS

A new category of high-risk MDS is defined according to p53 dysfunction

RNA sequencing of CD34+ progenitor cells form 236 MDS patients



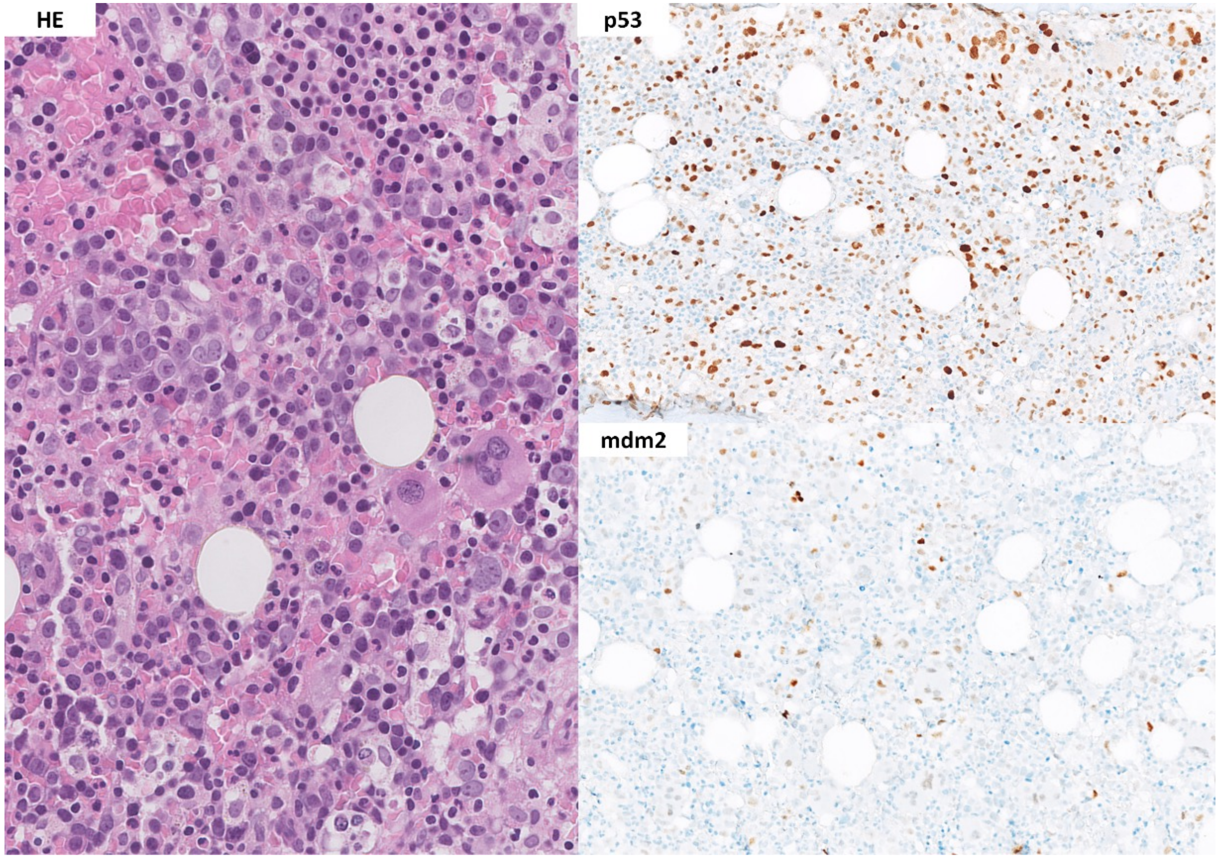
Evidence of impaired T cell and NK maturation and function in MDS with p53 dysfunction

- Immune checkpoint overexpression (PD-L1) at the stem cell level
- Reduced numbers of cytotoxic T cells
- Expansion of myeloid-derived suppressor cells (MDSCs)
- Expansion of regulatory T cells (Tregs).
- Impaired NK maturation and function

Sallman DA Blood (2020) 136 (24): 2812–2823

Riva et al. Blood 2022;140:4001–4; Zampini et al, manuscript in preparation

A new category of high-risk MDS is defined according to p53 dysfunction



GENOMED4ALL

Riva et al. Blood 2022;140:4001–4; Zampini et al, manuscript in preparation

2021 WHO guidance on ethics and governance of AI for health

We have to address three important topics, deemed as essential for a **right deployment of AI in hematology**:

- **Transparency of models.** We have to provide a good understanding of the models (interpretability and explainability)

- **Reliability of models.** The main vulnerabilities of AI models are related to lack of generalizability. Therefore, extensive, independent validation of generated AI-models is required.

- **Protection of data and data sharing**. Innovative technologies such as federated learning procedures for data collection and analysis (without moving sensitive medical data from their original locations) are required to facilitate clinical implementability of AI solutions

1. *The World Health Organization. 2021 WHO guidance on ethics and governance of artificial intelligence for health.* *https://www.who.int/publications/i/item/9789240029200*

Artificial Intelligenece (AI) for precision medicine

1- Machine Learning

2- Generative AI

HUMANITAS

Generation of Synthetic Data to accelerate Research & Development in MDS

- In MDS the first evidence of recurrent somatic mutations in splicing-related genes was published in 2011 and only 10 years later large patient populations (n>2,000) with comprehensive clinical and molecular information were available to test clinical significance and implementability of genomic screening

- The development of innovative data-driven digital health products and services, in fact, is currently slowed due to limited access to / availability of data. Additional challenges in health data include harmonization problems and data privacy (GDPR)

- **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 so they can be easily accessed and shared.

SYNTHEMA

Generation of synthetic data to accelerate Research & Development in MDS

- Synthetic data can be generated by using neural networks (Generative Adversarial Networks, GAN).

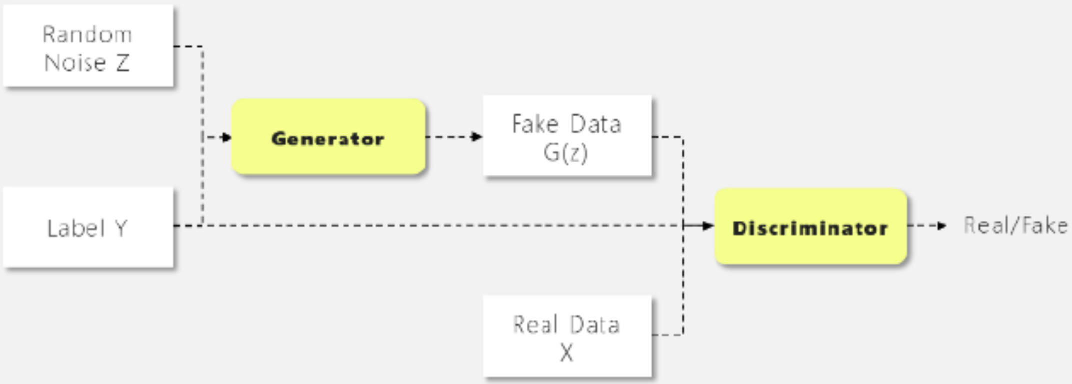


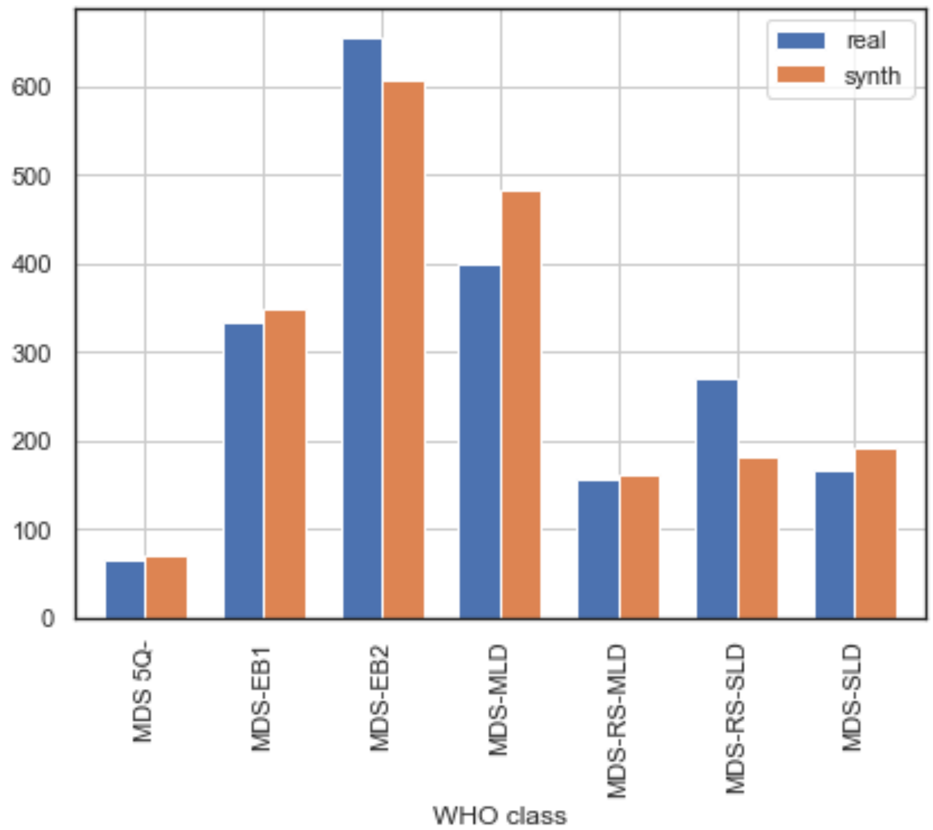
Figure 2 Conditional Generative Adversarial Networks architecture

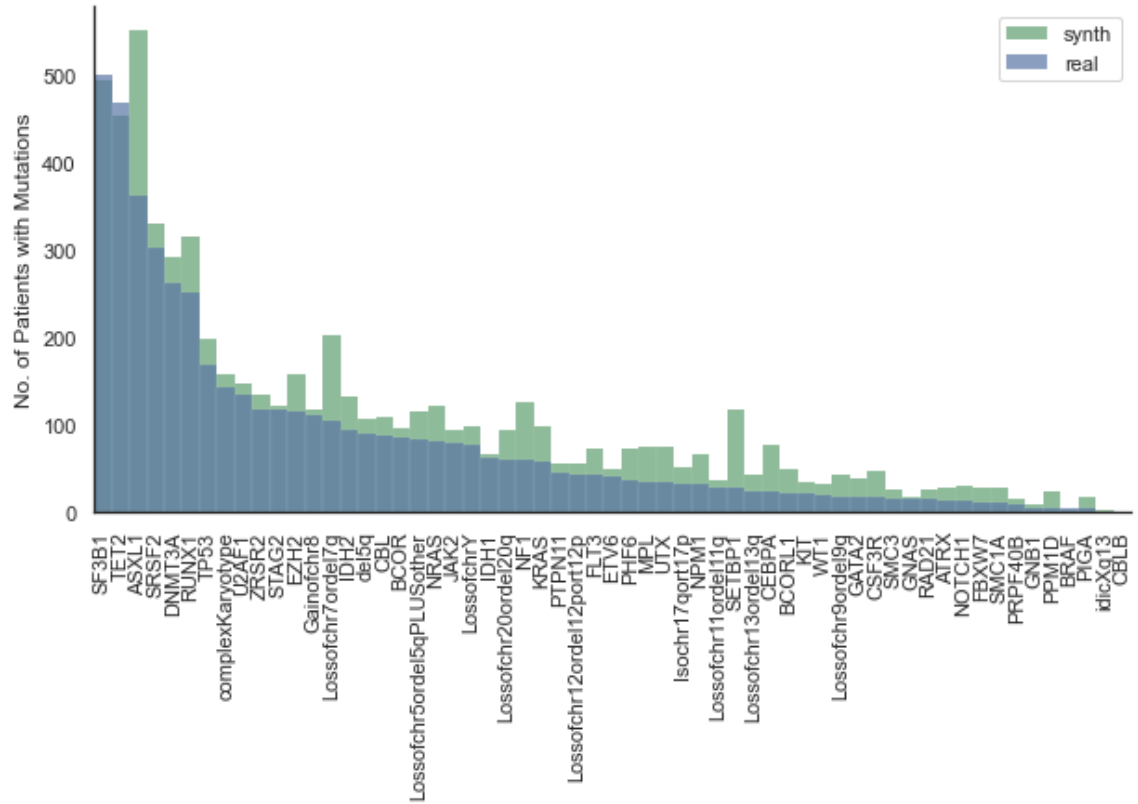
Possible applications

- Data sharing (GDPR)
- Classes balance and resolution of missing information
- Data augmentation for learning/ validation purpose
- Generation of new evidence

SYNTHEMA

Synthetic vs. Real Data: comparison of clinical and molecular features

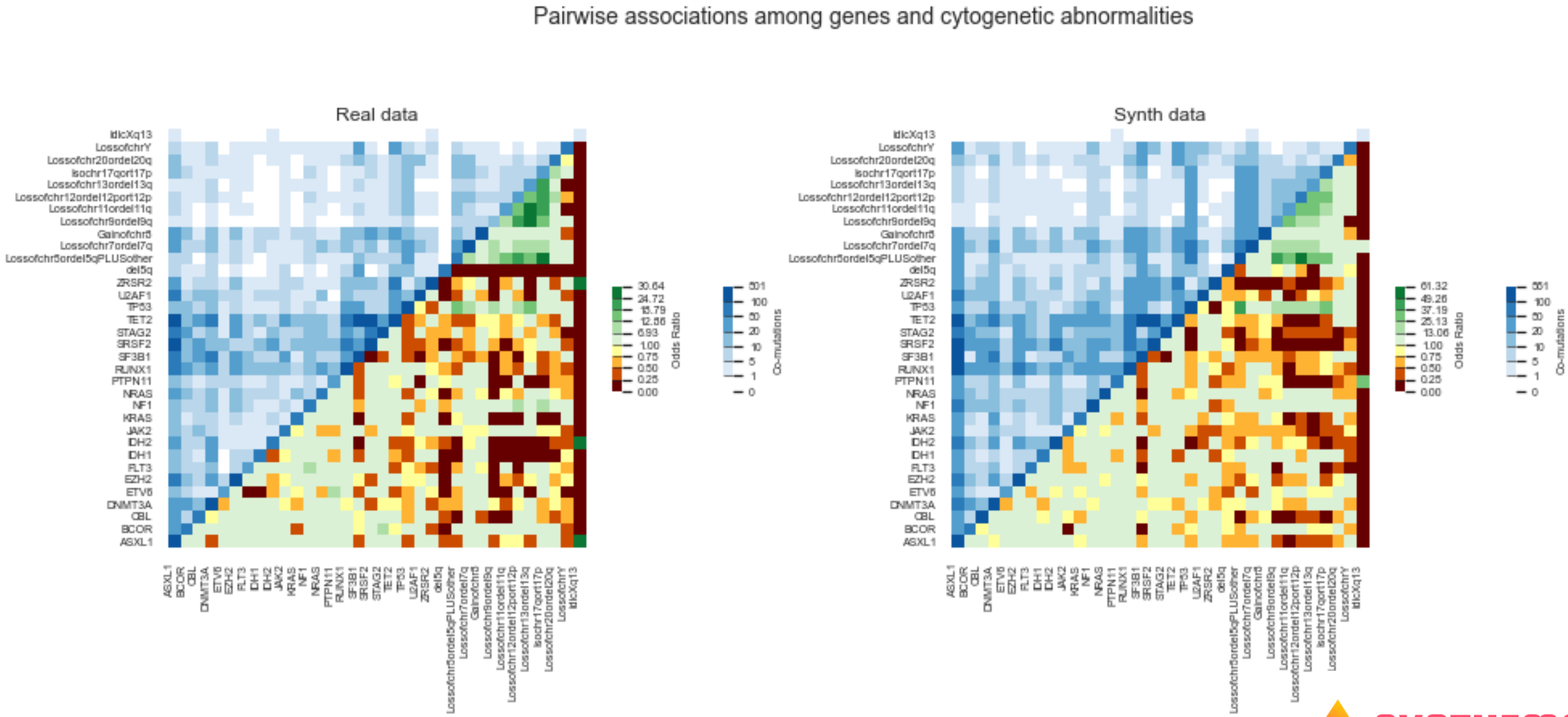




SYNTHEMA

D'Amico S et al, J Clin Oncol CCI 2023, in press

Synthetic vs. Real Data: pairwise association among genes and cytogenetic abnormalities



D'Amico S et al, J Clin Oncol CCI 2023, in press

SYNTHEMA

Synthetic vs. Real Data: survival

COX models (overall survival)

Real data:

Global Concordance: 0.779; Std.err:0.013

Partial Concordance of risk components:

	Clinical	CNA	Demographics	Genetics
concordant	0.711	0.569	0.630	0.782
std(c-d)	0.013	0.011	0.013	0.013

Synthetic data:

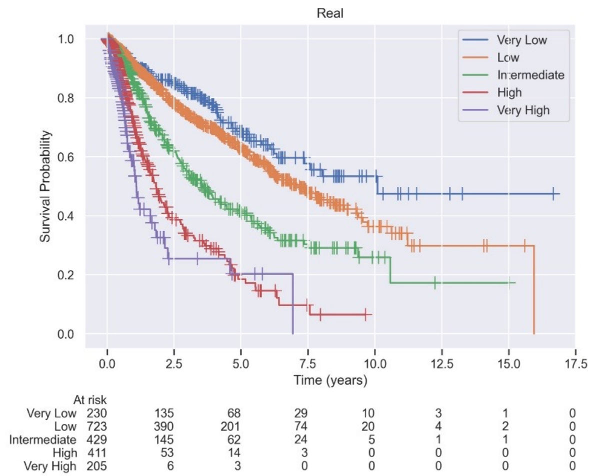
Global Concordance: 0.822; Std.err:0.013

Partial Concordance of risk components:

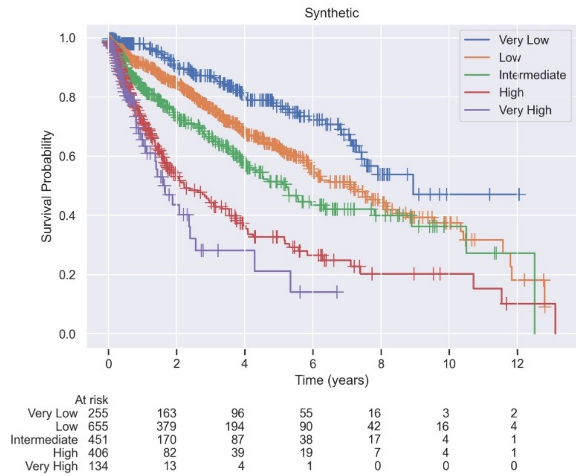
	Clinical	CNA	Demographics	Genetics
concordant	0.732	0.536	0.646	0.746
std(c-d)	0.013	0.011	0.013	0.013

Probability of OS stratified by IPSS-R

REAL DATA



SYNTHETIC DATA

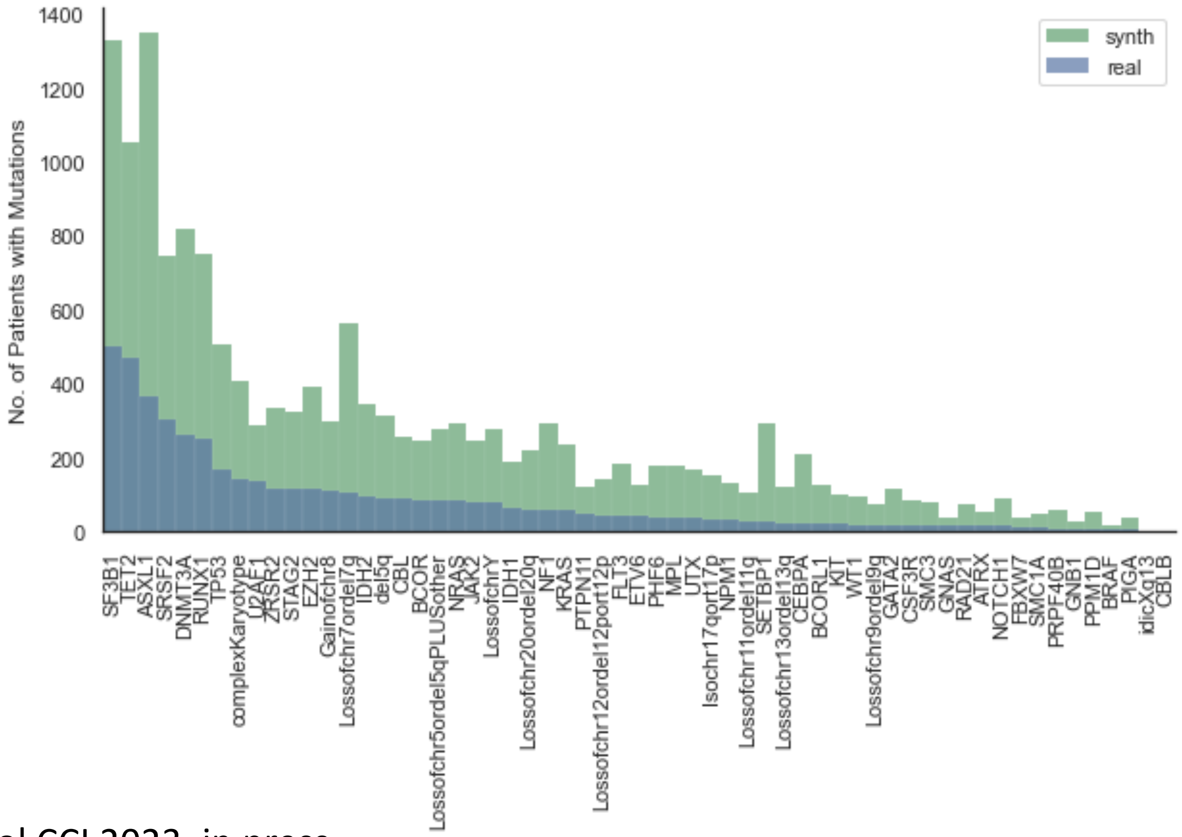


D'Amico S et al, J Clin Oncol CCI 2023, in press

SYNTHEMA

Synthetic vs. Real Data: data augmentation

Data augmentation: form 2043 to 5000 patients



D'Amico S et al, J Clin Oncol CCI 2023, in press

SYNTHEMA

Performance of Synthetic Data

DEMOGRAPHIC, CLINICAL AND SURVIVAL DATA

92.1 %

SYNTHETIC CLINICAL FITNESS

Evaluated with distribution plot, Principal Component Analysis and correlation matricies.

GENOMIC DATA

90.2 %

SYNTHETIC GENOMIC FITNESS

Evaluated with mutation frequencies and pairwise association.

ALL DATA

70.6 %

PRIVACY PRESERVABILITY

Evaluated considering the possibility of tracing real data from synthetic ones.

D'Amico S et al, J Clin Oncol CCI 2023, in press

SYNTHEMA

Generation of Synthetic Data to accelerate translational research in Hematology

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Somatic *SF3B1* Mutation in Myelodysplasia with Ring Sideroblasts

E. Papaemmanuil, M. Cazzola, J. Boultwood, L. Malcovati, P. Vyas, D. Bowen, A. Pellagatti, J.S. Wainscoat, E. Hellstrom-Lindberg, C. Gambacorti-Passerini, A.L. Godfrey, I. Rapado, A. Cvejic, R. Rance, C. McGee, P. Ellis, L.J. Mudie, P.J. Stephens, S. McLaren, C.E. Massie, P.S. Tarpey, I. Varela, S. Nik-Zainal, H.R. Davies, A. Shlien, D. Jones, K. Raine, J. Hinton, A.P. Butler, J.W. Teague, E.J. Baxter, J. Score, A. Galli, M.G. Della Porta, E. Travaglino, M. Groves, S. Tauro, N.C. Munshi, K.C. Anderson, A. El-Naggar, A. Fischer, V. Mustonen, A.J. Warren, N.C.P. Cross, A.R. Green, P.A. Futreal, M.R. Stratton, and P.J. Campbell for the Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium

6 patients
2011

NEJM Evidence

Published June 12, 2022
NEJM Evid 2022; 1 (7)
DOI: 10.1056/EVIDoa2200008

ORIGINAL ARTICLE

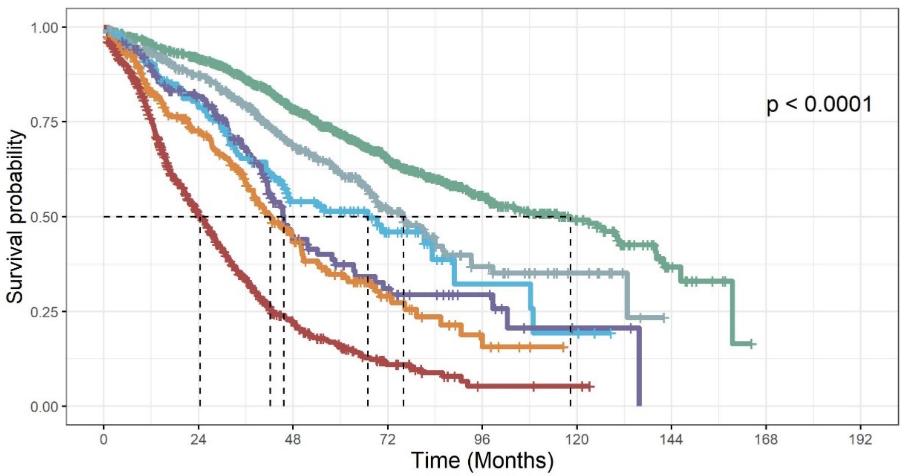
Molecular International Prognostic Scoring System for Myelodysplastic Syndromes

Elsa Bernard, Ph.D.,[1] Heinz Tuechler, Peter L. Greenberg, M.D.,[2] Robert P. Hasserjian, M.D.,[3] Juan E. Arango Ossa, M.S.,[1] Yasuhito Nannya, M.D., Ph.D.,[4,5] Sean M. Devlin, Ph.D.,[1] Maria Creignou, M.D.,[6] Philippe Pinel, M.S.,[1] Lily Monnier, M.S.,[1] Gunes Gundem, Ph.D.,[1] Juan S. Medina-Martinez, M.S.,[1] Dylan Domenico, B.S.,[1] Martin Jädersten, M.D., Ph.D.,[6] Ulrich Germing, M.D.,[7] Guillermo Sanz, M.D., Ph.D.,[8,9,10] Arjan A. van de Loosdrecht, M.D., Ph.D.,[11] Olivier Kosmider, M.D., Ph.D.,[12] Matilde Y. Follo, Ph.D.,[13] Felicitas Thol, M.D.,[14] Lurdes Zamora, Ph.D.,[15] Ronald F. Pinheiro, Ph.D.,[16] Andrea Pellagatti, Ph.D.,[17] Harold K. Elias, M.D.,[18] Detlef Haase, M.D., Ph.D.,[19] Christina Ganster, Ph.D.,[19] Lionel Ades, M.D., Ph.D.,[20] Magnus Tobiasson, M.D., Ph.D.,[6] Laura Palomo, Ph.D.,[21] Matteo Giovanni Della Porta, M.D.,[22] Akifumi Takaori-Kondo, M.D., Ph.D.,[23] Takayuki Ishikawa, M.D., Ph.D.,[24] Shigeru Chiba, M.D., Ph.D.,[25] Senji Kasahara, M.D., Ph.D.,[26] Yasushi Miyazaki, M.D., Ph.D.,[27] Agnes Viale, Ph.D.,[28] Kety Huberman, B.S.,[28] Pierre Fenaux, M.D., Ph.D.,[20] Monika Belickova, Ph.D.,[29] Michael R. Savona, M.D.,[30] Virginia M. Klimek, M.D.,[18] Fabio P. S. Santos, M.D., Ph.D.,[31] Jacqueline Boultwood, Ph.D.,[17] Ioannis Kotsianidis, M.D., Ph.D.,[32] Valeria Santini, M.D.,[33] Francesc Solé, Ph.D.,[21] Uwe Platzbecker, M.D.,[34] Michael Heuser, M.D.,[14] Peter Valent, M.D.,[35,36] Kazuma Ohyashiki, M.D., Ph.D.,[37] Carlo Finelli, M.D.,[38] Maria Teresa Voso, M.D.,[39] Lee-Yung Shih, M.S.,[40] Michaela Fontenay, M.D., Ph.D.,[12] Joop H. Jansen, Ph.D.,[41] José Cervera, M.D., Ph.D.,[42] Norbert Gattermann, M.D.,[7] Benjamin L Ebert, M.D., Ph.D.,[43] Rafael Bejar, M.D., Ph.D.,[44] Luca Malcovati, M.D.,[45] Mario Cazzola, M.D.,[45] Seishi Ogawa, M.D., Ph.D.,[4,46,47] Eva Hellström-Lindberg, M.D., Ph.D.,[6] and Elli Papaemmanuil, Ph.D.[1]

2957 patients
2022

Generation of Synthetic Data to accelerate translational research in Hematology

IPSS-M (real data, 2022)

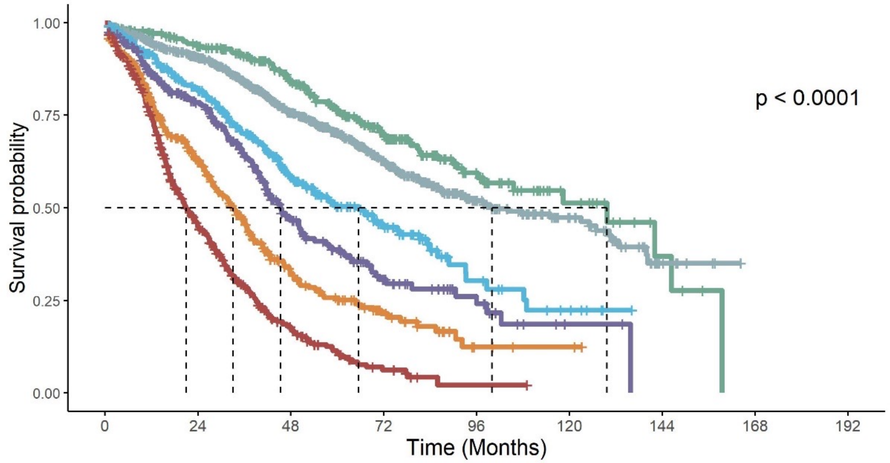


No. at risk

IPSS-M	0	24	48	72	96	120	144	168	192
	1074	884	587	317	125	53	11	0	0
	379	267	155	70	23	9	0	0	0
	149	101	45	22	5	3	0	0	0
	155	100	36	20	9	2	0	0	0
	236	133	58	19	6	0	0	0	0
	746	308	91	21	3	2	0	0	0

Time (Months)

Syntetic IPSS-M (synthetic data, 2013)



No. at risk

Synthetic IPSS-M	0	24	48	72	96	120	144	168	192
	312	266	193	109	45	15	4	0	0
	964	768	483	250	94	44	7	0	0
	326	218	111	50	14	6	0	0	0
	286	172	72	30	13	2	0	0	0
	381	192	64	22	4	2	0	0	0
	470	177	49	8	1	0	0	0	0

Time (Months)

D'Amico S et al, J Clin Oncol CCI 2023, in press

SYNTHEMA

Generation of Synthetic Data to accelerate clinical research in Hematology

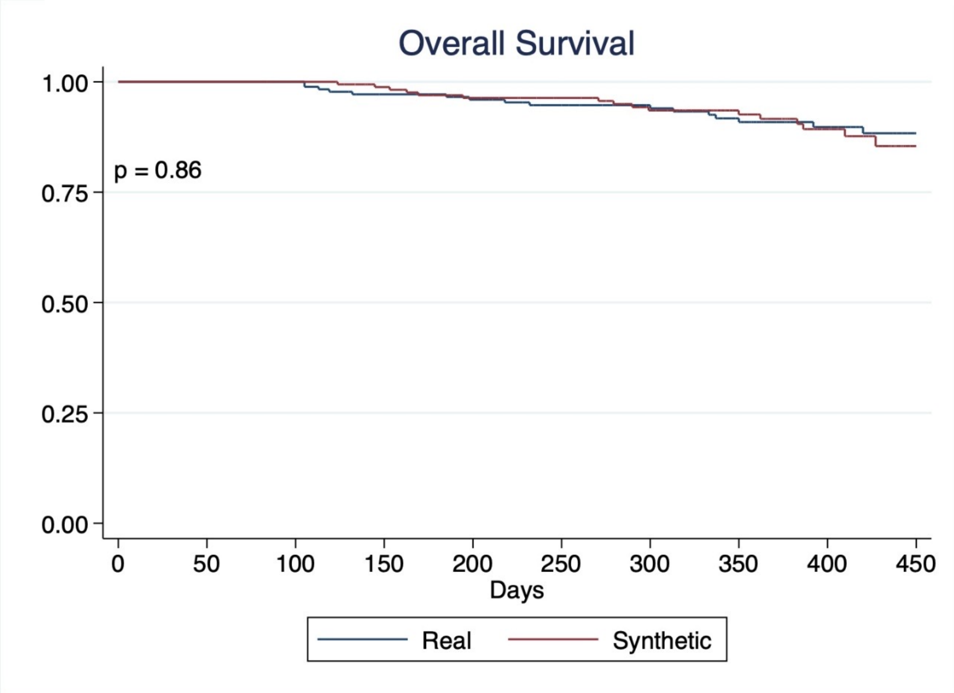
Comparing endpoints of clinical trials using **real** and **synthetic** control arms. Real-world efficacy and safety of luspatercept in adult patients with transfusion-dependent anemia due to very low-, low and intermediate-risk myelodysplastic syndrome (MDS) with ring sideroblasts, who had an unsatisfactory response to or are ineligible for erythropoietin-based therapy: a multicenter study by Fondazione Italiana Sindromi Mielodisplastiche (FISIM)

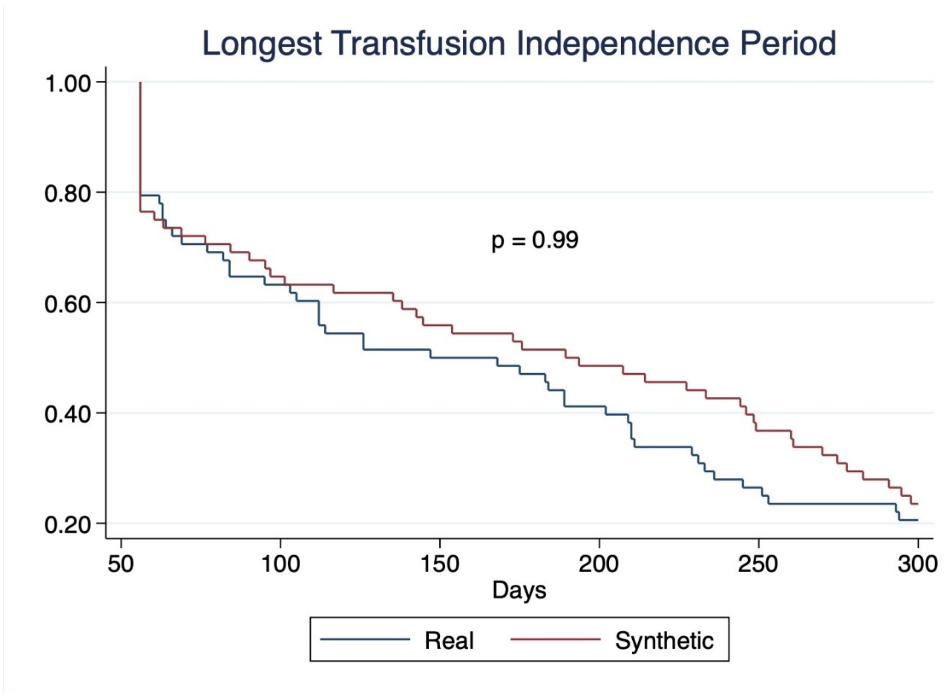
Primary endpoint

	Real data	Synthetic data	Pvalue
RBC-TI>=8 weeks 1-24	56 (31,5)	56 (31,5)	1.0
Longest Transfusion Independence Period (weeks), median (range)	195 (56-490)	191 (56-490)	0.34
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 RBC	62 (34,8)	63 (35,4)	1.0
Reduction>=50%	77 (43,3)	72 (40,4)	0.66
AML Evolution	4 (2,2)	6 (3,4)	0.75
Discontinued patients	74 (41,6)	82 (46,1)	0.64

D'Amico S et al, J Clin Oncol CCI 2023, in press

Generation of Synthetic Data 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 adult patients with transfusion-dependent anemia due to very low-, low and intermediate-risk myelodysplastic syndrome (MDS) with ring sideroblasts, who had an unsatisfactory response to or are ineligible for erythropoietin-based therapy: a multicenter study by Fondazione Italiana Sindromi Mielodisplastiche (FISIM)

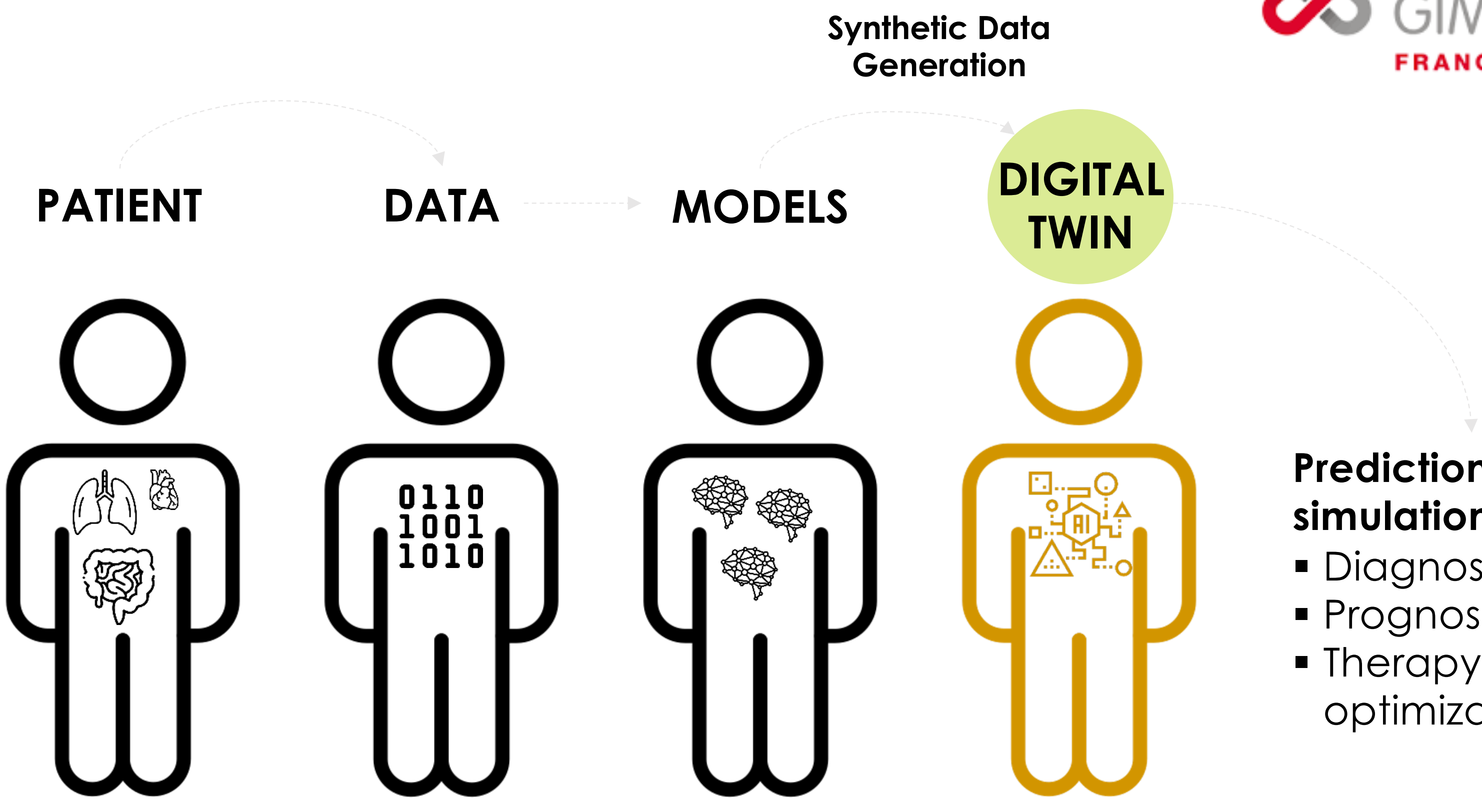
D'Amico S et al, J Clin Oncol CCI 2023, in press

SYNTHEMA

From Synthetic data to Digital Twins

fondazione GIMEMA onlus
FRANCO MANDELLI

Different data layers:
- Clinical
- Genomic
- Images



Prediction and simulation:
- Diagnosis
- Prognosis
- Therapy optimization

HUMANITAS

C.A.L.R.

CENTER FOR ACCELERATING LEUKEMIA/LYMPHOMA RESEARCH

Artificial Intelligence and real world data analysis to improve patient care and advance medical research in hematology

HUMANITAS HU HUMANITAS UNIVERSITY POLITECNICO MILANO 1863

SYNTHEMA

GENOMED4ALL

HU HUMANITAS UNIVERSITY Horizon2020 European Union Funding for Research & Innovation HUMANITAS RESEARCH HOSPITAL

AI people

- Saverio D'Amico
- Elisabetta Sauta
- Gianluca Asti
- Victor Savevski
- Gastone Castellani

Clinical team

- Luca Lanino
- Giulia Maggioni
- Erica Travaglino
- Alessia Campagna
- Marta Ubezio
- Antonio Russo
- Gabriele Todisco
- Armando Santoro