

P495 UNLOCKING THE POTENTIAL OF SYNTHETIC PATIENTS FOR ACCELERATING CLINICAL TRIALS: RESULTS OF THE FIRST GIMEMA EXPERIENCE

Topic: 4. Acute myeloid leukemia - Clinical

Alfonso Piciocchi¹, Marta Cipriani^{2,1}, Monica Messina¹, Giovanni Marconi³, Valentina Arena¹, Stefano Soddu¹, Enrico Crea¹, Maria Valeria Feraco⁴, Marco Ferrante⁴, Edoardo La Sala¹, Paola Fazi¹, Francesco Buccisano⁵, Giovanni Martinelli⁶, Adriano Venditti⁵, Marco Vignetti¹

¹Gimema Foundation, Rome, Italy; ²Sapienza University, Rome, Italy; ³Hematology Unit, Irccs Istituto Romagnolo Per Lo Studio Dei Tumori (Irst) "dino Amadori", Meldola, Italy; ⁴Studio Legale Flc, Rome, Italy; ⁵Division Of Hematology, Tor Vergata Foundation Polyclinic, Rome, Italy; Hematology, Department Of Biomedicine E Prevention, Tor Vergata University, Rome, Italy; ⁶Irccs Istituto Romagnolo Per Lo Studio Dei Tumori (Irst) "dino Amadori", Rome, Italy

Background:

Artificial intelligence is contributing to improve different medicine areas including clinical trial design. One field that holds a great potential is represented by the use of digital data as an alternative to real ones. The generation of a virtual cohort of patients might be advantageous since an artificial group of patients resembles the real dataset in all the key features but it does not include any identifiable real-patient data, tackling - by a privacy standpoint - the "burden" of collecting data subjects' consent as well as the shortcomings of common anonymization techniques.

Aims:

To test the feasibility of this approach and evaluate its potential in clinical trial design, we built an in-silico cohort based on the large dataset of patients enrolled in the GIMEMA AML1310 study (Venditti et al. 2019), which entailed a "3+7"-like induction and a risk-adapted, MRD-directed post-remission transplant allocation.

Methods:

To create the synthetic cohort of patients, a machine learning generative model was constructed from the real individual-level data of the AML1310 study, capturing its patterns and statistical properties. AML1310 enrolled 500 patients (median age 49 years old) in 55 GIMEMA Institutions. All patients were NCCN2009 risk classified and analyzed by morphology, cytogenetics, molecular biology and multiparametric flow cytometry. The subset of 445 patients with ELN2017 risk classification available was used.

To this purpose, the R package "synthpop" was used considering a parametric method: for binary data the logistic regression, for a factor with > 2 levels the polytomous logistic regression, for an ordered factor with > 2 levels the ordered polytomous logistic regression. For time to event variables the classification and regression trees method was used.

Next, we verified the adherence of the virtual cohort to the original one in terms of age, gender, PS, WBC count, FLT3 and NPM1 mutations, risk category, CR achievement, MRD, transplant rate. Virtual and real cohorts were also compared in terms of survival outcomes.

Results:

By using the real-patient dataset from the AML1310 trial, a virtual cohort of 850 patients, named synthAML1310, was generated. By comparing the two cohorts, we observed that the clinico-biological characteristics and response evaluations (CR and MRD rates) did not differ significantly. Moreover, as depicted in Figure 1, the curves of OS

Copyright Information: (Online) ISSN: 2572-9241

© 2023 the Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the European Hematology Association. This is an open access Abstract Book distributed under the Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) which allows third parties to download the articles and share them with others as long as they credit the author and the Abstract Book, but they cannot change the content in any way or use them commercially.

Abstract Book Citations: Authors, Title, HemaSphere, 2023;7(S3):pages. The individual abstract DOIs can be found at <https://journals.lww.com/hemasphere/pages/default.aspx>.

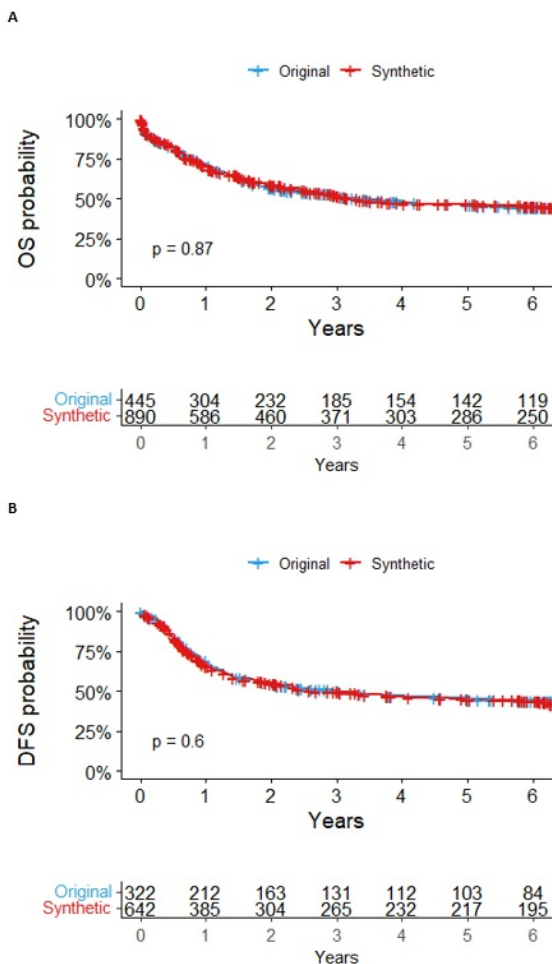
Disclaimer: Articles published in the journal HemaSphere exclusively reflect the opinions of the authors. The authors are responsible for all content in their abstracts including accuracy of the facts, statements, citing resources, etc.

and DFS were superimposable. Indeed, at 2 years, OS was 57% (52.5%-61.9%) in the original and 59.1% (55.9%-62.6%) in the synthAML1310 cohort. DFS was 55.1% (49.8%-60.9%) in the original and 55.1% (51.3%-59.2%) in the synthetic cohort.

Summary/Conclusion:

These results demonstrate the success of this approach in producing a virtual dataset that perfectly mimics the original and that, from a “privacy by design” perspective, minimizes the risk of re-identification of patients. Mirroring an AML population treated with a conventional chemotherapeutic approach, synthAML1310 is suitable to represent the control group when testing novel innovative treatments, most likely in an in-silico randomized trial, but also in other settings like propensity score matching analyses in observational studies. Shifting to an in-silico trial would overcome the challenges of randomized trials and it would be beneficial also for patients. since, they would receive only the experimental treatment without being exposed to the “less active” therapy, thus limiting treatment failures and toxicity. Also, enrolment and the attainment of final results would be faster.

Figure 1. Survival outcomes of the original and synthetic AML1310 cohorts. A) Overall Survival and B) Disease-free survival.



Copyright Information: (Online) ISSN: 2572-9241

© 2023 the Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the European Hematology Association. This is an open access Abstract Book distributed under the Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) which allows third parties to download the articles and share them with others as long as they credit the author and the Abstract Book, but they cannot change the content in any way or use them commercially.

Abstract Book Citations: Authors, Title, HemaSphere, 2023;7(S3):pages. The individual abstract DOIs can be found at <https://journals.lww.com/hemasphere/pages/default.aspx>.

Disclaimer: Articles published in the journal HemaSphere exclusively reflect the opinions of the authors. The authors are responsible for all content in their abstracts including accuracy of the facts, statements, citing resources, etc.

Copyright Information: (Online) ISSN: 2572-9241

© 2023 the Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the European Hematology Association. This is an open access Abstract Book distributed under the Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) which allows third parties to download the articles and share them with others as long as they credit the author and the Abstract Book, but they cannot change the content in any way or use them commercially.

Abstract Book Citations: Authors, Title, HemaSphere, 2023;7(S3):pages. The individual abstract DOIs can be found at <https://journals.lww.com/hemasphere/pages/default.aspx>.

Disclaimer: Articles published in the journal HemaSphere exclusively reflect the opinions of the authors. The authors are responsible for all content in their abstracts including accuracy of the facts, statements, citing resources, etc.