PH-0072

MACHINE LEARNING BASED PREDICTIVE MODELING FOR MORTALITY 100 DAYS POST ALLOGENEIC HEMATOPOIETIC STEM CELL TRANSPLANTATION (ALLO-HSCT) IN ACUTE LEUKEMIA: AN EBMT- ACUTE LEUKEMIA WORKING PARTY (ALWP) REGISTRY STUDY

R. Shouval[1,], M. Labopin[2], O. Bondi[3], H. Mishan-Shamay[3], A. Shimoni[4], F. Ciceri[5], J. Esteve[6], S. Giebel[7], N.-C. Gorin[2], C. Schmid[8], I. Zakaria[2], L. Moukhtari[2], E. Polge[2], M. Aljurf[9], N. Kroger[10], C. Craddock[11], A. Bacigalupo[12], J. Cornelissen[13], F. Baron[14], R. Unger[3], A. Nagler [4], M. Mohty[2] and Acute Leukemia Working Party of the EBMT*

[1]Hematology and Bone Marrow Transplantation, Chaim Sheba Medical Center, Tel-Hashomer, Ramat- Gan, Israel, [2]EBMT Acute Leukemia Working Party and Registry, Department of Hematology and Cell Therapy, Hospital Saint-Antoine, Paris University, Paris, France, [3]The Mina and Everard Goodman Faculty of Life Sciences, Department of Computational Biology and Bioinfromatics, Bar-Ilan University, [4]Hematology and Bone Marrow Transplantation, Chaim Sheba Medical Center, Tel-Hashomer, Ramat Gan, Israel, [5]Hematology and BMT Unit, San Raffaele Scientific Institute, Milan, Italy, [6]Hematology, IDIBAPS, Hospital Clinic, Barcelona, Spain, [7]Bone Marrow Transplantation, Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Gliwice, Poland, [8]Hematology and Oncology, Klinikum Augsburg,Ludwig-Maximilinas-Universitat, Munich, Germany, [9]King Faisal Specialist Hospital and Research Centre - Oncology Center, Riyadh, Saudi Arabia, [10]Center of Oncology-Department of Stem Cell Transplantation, UNIVERSITY MEDICAL CENTER Hamburg Eppendorf, Hamburg, Germany, [11]Centre for Clinical Haematology, Dept of Haematology, University of Birmingham and Queen Elizabeth Hospital, Birmingham, United Kingdom, [12]Department of Hematology Oncology, Ospedale San Martino, Genoa, Italy, [13]Hematology, Erasmus MC, Rotterdam, Netherlands, [14]Hematology & GIGA research, University of Liège, Liege, Belgium

Introduction: Machine learning is a field in computer science stemming from artificial intelligence and is part of the data mining approach for data analysis. It is commonly applied in financial and technological settings, where data scenarios are complex. We hypothized that given the complexity of allo-HSCT patient data, applying a data mining approach, may yield improved outcome prediction models, as compared to current risk score, which are based on a conventional statistical approach. Our aim was to develop prediction models for 100 days post allo-HSCT overall mortality (OM) and non-relapse related mortality (NRM).

Materials (or patients) and Methods: A cohort of 28,236 adult allo-HSCT recipients from the registry of the ALWP of the EBMT was analyzed. Twenty variables were analyzed, including year of transplant (range, 2000-2011), diagnosis (Acute Myeloid Leukemia and Acute Lymphoblastic Leukemia), disease stage (CR1, CR2, advanced/ refractory disease), conditioning regimen (myeloablative or reduced- intensity conditioning), graft type (peripheral blood or bone marrow), donor type (HLA matched related or unrelated), etc. The dataset was randomly divided into a training set (n=19,765) and a validation set (n=8,707). An alternating decision tree (ADT) machine learning algorithm was trained and optimized on the former and validated on the last. The ADT prediction models for OM and NRM at day 100 were assessed according to (a) area under the curve (AUC) of the receiver operating characteristic curve (b) variable inclusion (c) calibration. Additionally, a standard univariate analysis was applied.

Results: The ADT prediction models' AUCs for OM and NRM were 0.70 (95% [CI] 0.69–0.70) and 0.67 (95% [CI] 0.66–0.67) respectively. Ten mutual variables were selected by the algorithm, however, weights and dependence status varied between models. Calibration between models performance on training and validation sets was excellent (R^2=0.9895 and R^2=0.9627 respectively). Probability for 100 day OM and NRM ranged from 0.06-0.52 and 0.04–0.39 respectively. For each model, patients were stratified into 6 risk groups.

Discussion: We present 2 novel prediction models for OM and NRM at day 100 post allo-HSCT. The models enable risk assessment, in a stratified manner, prior to transplantation. Data from approximately 30,000 patients was used for model development, assuring robustness and validity. Incorporation of our model with conventional risk scores, such as the Hematopoietic Cell Transplant-Co-morbidity Index, may further improve predictive performance.

In summary, the alternating decision tree prediction models may aid outcome prediction and risk evaluation in patients with acute leukemia prior to allo-HSCT.

Disclosure of Interest: None Declared.