Predictive limitations of hematopoietic stem cell transplantation associated mortality: a machine learning in-silico analysis of the EBMT - Acute Leukemia Working Party Registry

Roni Shouval, Myriam Labopin, Ron Unger, Sebastian Giebel, Fabio Ciceri, Christoph Schmid, Jordi Esteeve, Norbert Gorin, Frédéric Baron, Bipin N. Savani, Mohamad Mohty, Arnon Nagler

Angaben zur Veröffentlichung / Publication details:

Shouval, Roni, Myriam Labopin, Ron Unger, Sebastian Giebel, Fabio Ciceri, Christoph Schmid, Jordi Esteeve, et al. 2015. "Predictive limitations of hematopoietic stem cell transplantation associated mortality: a machine learning in-silico analysis of the EBMT - Acute Leukemia Working Party Registry." *Biology of Blood and Marrow Transplantation* 21 (2): S310–11. https://doi.org/10.1016/j.bbmt.2014.11.495.

Conclusion: In conclusion, although a similar rate of attrition was seen in donors overall, the reasons were different from our previous study. Two of the changes instituted are likely to have contributed (lowering joining age and increasing BMI). Further work must be done on understanding other factors associated with attrition, and collaborating with other international registries to permit access to donors who have moved to other countries.

448

Differential Impact of Dose Escalated Busulfan on Allogeneic Transplant for High, Intermediate and Low Risk Disease

Thomas C. Shea [1], Christine M. Walko [2], Yunro Chung [3], Anastasia Ivanova [4], Kamakshi V. Rao [5], James Coghill [6], Stefanie Sarantopoulos [7], William A. Wood [8], Paul Armistead [9], Don A. Gabriel [10], Jonathan S. Serody [11].
[1] *University of North Carolina at Chapel Hill, Chapel Hill, NC;* [2] *Department of Pharmacology, Moffitt Cancer Center, Tampa, FL;* [3] *Biostatistics, UNC School of Public Health, Chapel Hill, NC;* [4] *Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC;* [5] *Department of Pharmacy, University of North Carolina Hospitals and Clinics, Chapel Hill, NC;* [6] *Univ of North Carolina, Chapel Hill, NC;* [7] *Duke Adult Blood and Marrow Transplant Program, Duke University, Durham, NC;* [8] *Division of Hematology/Oncology, University of North Carolina - Chapel Hill, Chapel Hill, NC;* [9] *Hematology/ Oncology, University of North Carolina, Chapel Hill, NC;* [10] *Univ of NC School of Medicine, Chapel Hill, NC;* [11] *Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC*

Disease relapse, graft vs host disease and infection remain the major barriers to successful allogeneic stem cell transplantation. We previously presented data on the use of escalated AUC based dosing of a continuous infusion (CI) of IV busulfan over four days (Walko, BBMT, S218, 2012). In that report, we described use of a test dose and day 1 and 4 PK values to determine the dose delivered, and identified dose limiting toxicities as rash and mucositis at an AUC of 8300 um-min/day and identified an AUC MTD of approximately 6912 uM-min/day x 4 days; a 40% increase over a standard AUC dose of 4800 uM-min/day. Here we report additional analyses to identify patients (pts) likely to benefit from this higher dose of continuous infusion busulfan.

Methods: Patients with advanced hematologic malignancies and adequate organ function who were appropriate for a MUD or MRD allogeneic transplant were enrolled on an IRB approved trial of a test dose of busulfan (.8 mg/kg x 1) followed by escalated dose busulfan (4800-8300 uM-min/d) given as a 90 hour continuous infusion along with fludarabine at 30 mg/m^2/d x 5. Pts received tacrolimus and either alemtuzamab (30), ATG + MTX (19) or MTX alone (6) as GVH prophylaxis with standard anti-infective and supportive care.

Results: 55 pts (median age 41, range 20-55) with myeloid (38), lymphoid (8), or biphenotypic (1) leukemias, MDS or MF (1) or lymphomas (7) were enrolled. All 55 subjects were analyzed according to their being in the AUC low (group 1, 19 pts), middle (group 2, 19 pts), or high dose (group 3, 17 pts). 17 subjects had low, 20 had intermediate, and 18 had high-risk disease by CIBMTR criteria. For the entire group, univariate analysis identified age, recipient CMV status, and disease risk as significant factors for overall (OS) and relapse free survival (RFS). Co-morbidity scores, GVH occurrence or prophylaxis, donor/recipient sex and donor type (MUD or MRD) were not significant, nor was the non-relapse mortality rate different between the three AUC groups. Multivariable analysis identified CMV status as borderline significant (p=.07), and high vs low AUC dose (p=.053) and disease risk (p=.01) as significant for both OS and RFS. Outcomes were similar between AUC groups 2 and 3. Differences in OS and RFS were limited to the good and intermediate risk patients as outcomes for high-risk patients were poor for all AUC groups (0/17 RFS and 1/17 OS). When analyzing the 38 good and intermediate risk patients, the OS and RFS were 66% and 62.5% for combined AUC groups 2 and 3 compared to 31% (p=.02) for both OS and RFS in AUC group 1.

Conclusions: Targeted, PK based, CI busulfan that is approximately 40% higher than standard doses can result in apparent benefit for patients with CIBMTR low or intermediate risk disease. These doses were unable to demonstrate benefit in high-risk patients for whom other approaches such as immunotherapy, hypomethylating agents, or small molecule inhibitors may be needed.

449

Predictive Limitations of Hematopoietic Stem Cell Transplantation Associated Mortality: A Machine Learning in-Silico Analysis of the EBMT - Acute Leukemia Working Party Registry

Roni Shouval [1], Myriam Labopin [2], Ron Unger [3], Sebastian Giebel [4], Fabio Ciceri [5], Christoph Schmid [6], Jordi Esteeve [7], Norbert Gorin [8], Frédéric Baron [9], Bipin N. Savani [10], Mohamad Mohty [11,12], Arnon Nagler [1,12].
[1] *The Chaim Sheba Medical Center, Tel-Hashomer, Division of Hematology and Bone Marrow Transplantation, Ramat-Gan, Israel;* [2] *EBMT Paris study office / CEREST-TC, Paris, France;* [3] *The Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan, Israel;* [4] *Department of Bone Marrow Transplantation and Oncohematology, Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Gliwice Branch, Gliwice, Poland;* [5] *Hematology and Bone Marrow Transplantation Unit, San Raffaele Scientific Institute, Milan, Italy;* [6] *Department of Hematology and Oncology, Klinikum Augsburg, Ludwig-Maximilians-University, Munich, Germany;* [7] *Hematology Department, IDIBAPS, Hospital Clínic, Barcelona, Spain;* [8] *Hopital Saint-Antoine, Paris, France;* [9] *University of Liège, GIGA-I3, Liège, Belgium;* [10] *Medicine, Vanderbilt University, Brentwood, TN;* [11] *Department of Haematology, Saint Antoine Hospital, Paris, France;* [12] *EBMT Paris Office, Hospital Saint Antoine, Paris, France*

Several risk scores have been developed for the prediction of transplant related mortality (TRM) following allogeneic hematopoietic stem cell transplantation (HSCT). These have been validated; however, predictive performance is suboptimal. In addition to inherent uncertainty in such a complex medical procedure, methodological factors impeding prediction might be attributed to the statistical methodology, number and quality of features collected, or simply the population size. Using an *in-silico* approach (i.e. iterative computerized simulations), based on machine learning (ML) algorithms, we set to explore the factors limiting prediction.

ML is a subfield of computer science and artificial intelligence that deals with the construction and study of systems that can learn from data, rather than follow explicitly programmed instructions. Commonly applied in complex data scenarios, such as financial and technological settings, it may be suitable for outcome prediction of HSCT.

Study design involved two phases. The first, focused on development of several ML based prediction models of day

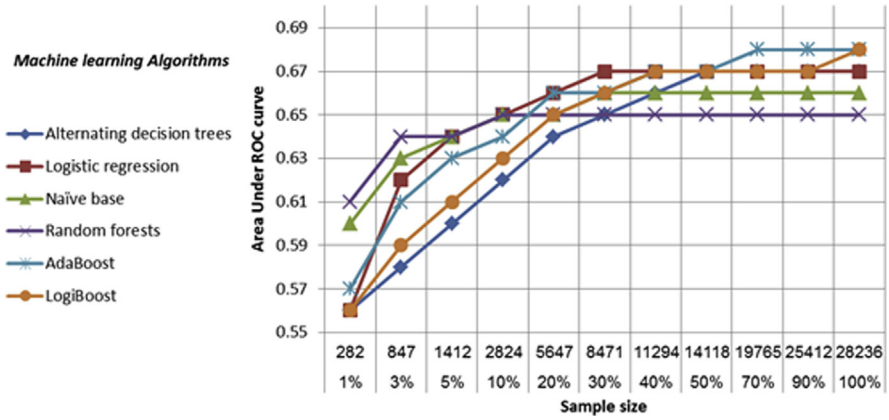

Figure 1. Predictive performance of TRM prediction models, adjusted for small size

100 TRM. A cohort of 28,236 acute leukemia, adult allogeneic HSCT recipients were analyzed. Twenty four variables were included. In the second phase, by applying a repetitive computerized simulation, factors necessary for optimal prediction were explored: algorithm type, size of data set, number of included variables, and performance in specific subpopulations. Models were assessed and compared on the basis of the area under the receiver operating characteristic curve (AUC).

We developed 6 ML based prediction models for day 100 TRM. Optimal AUCs ranged from 0.65-0.68. Predictive performance plateaued for a population size ranging from n=5647-8471, depending on the algorithm (Figure 1). A feature selection algorithm ranked variables according to importance. Provided with the ranked variable we data, discovered that a range of 6-12 ranked variables were necessary for optimal prediction, depending on the algorithm. Predictive performance of models developed for specific subpopulations ranged from an average of 0.59 to 0.67 for patient in second complete remission and patients receiving reduced intensity conditioning respectively.

In summary, we present a novel computational approach for prediction model development and analysis in the field of HSCT. Using data commonly collected on transplant patients, our simulation elucidates outcome prediction limiting factors. Regardless of the methodology applied, predictive performance converged when sampling more than 5000 patients. Few variables "carry the weight" with regard to predictive influence. Overall, the presented findings reveal a phenomenon of predictive saturation with data traditionally collected. Improving predictive performance will likely require additional types of input like genetic, biologic and procedural factors.

450

Discovery Analysis of Associations Between MicroRNAs (MiRs) and Both Pre-Transplant Comorbidity Burden and Post-Transplant Mortality in Patients (Pts) with Acute Leukemia (AL) in Complete Remission (CR) Given Allogeneic Hematopoietic Cell Transplantation

Mohamed L. Sorror [1,2], Kirsteen Maclean [3], Shamali Roy [3], A. Mario Marcondes [1,2], Muneesh Tewari [4], Beverly Torok-Storb [1,2]. [1] *Fred Hutchinson Cancer Research Center, Seattle, WA;* [2] *University of Washington, Seattle, WA;* [3] *NanoString Technologies, Seattle, WA;* [4] *University of Michigan, Ann Arbor, MI*

The HCT-comorbidity index (CI) was developed as a measure of health-status that could stratify risks of mortality after HCT. There is a great need for novel biomarkers that could explain the biologic link but also increase the objectivity of diagnosing pre-HCT comorbidities and increase the predictive power for post-HCT mortality. MiRs are a class of small non-coding RNAs (~22 nt) that negatively regulate gene expression. Studies have uncovered the functional role of miRs in diverse pathophysiological processes. Moreover, a single miR could be implicated in different pathological processes. To this end, we analyzed miRs as diagnostic for comorbidities before and as prognostic for mortality after HCT.

Peripheral blood mononuclear cell samples were previously collected from 36 pts within 30 days prior to HCT as a part of research repository. All samples were collected in EDTA tubes and processed and frozen at -80 Celsius degrees within 8 hours of draw. All pts were in CR before HCT. Low risk was defined as having HCTCI score of 0 before and surviving after HCT (median follow up 56 month, range 12.5-75.5), while high-risk pts had scores of 4-9 before HCT and none of them survived HCT (Table 1).

RNA was isolated from PBMC using previously described methods (Xie LN et al, Clinical Transplant. 2014; 28:314). For discovery of relevant miRs, we used NanoString nCounter miR assay as previously described (Knouf EC et al, 2013. PLoS ONE 8: e69630) comprising 654 endogenous miRs. Analysis of miR raw data was done using nSolver™ 2.0 Software (NanoString Technologies, Inc.) applying standard quality control tests. All samples contributed to the discovery analysis. MiRs

Table 1

Characteristics		%	
		Low risk (n=18)	High risk (n=18)
Donor	Related	17	28
	Unrelated	83	72
AL	Myeloid	78	67
	Lymphoid	22	33
CR#	1st	67	61
	2nd	33	39
Conditioning	High-dose	50	39
	Reduced-intensity	50	61