The 65th ASH Annual Meeting Abstracts

ONLINE PUBLICATION ONLY

632.CHRONIC MYELOID LEUKEMIA: CLINICAL AND EPIDEMIOLOGICAL

Identification of Individual Risk Factors for Treatment Free Remission (TFR) Failure in Chronic Phase Chronic Myeloid Leukemia (CP-CML) Patients: Predictive Models Developed on a Nationwide Clinical and Claim Database

Jules Chassetuillier, DipEng [1], *Meriem Benmerad, MSc, DipEng* [1], *Pierre-Louis Lecoq, MSc* [2], *Romain Favier, PharmD* [2], *Matthieu Trancart, MSc, Eng* [2], *Sandrine Hayette, PhD* [3,4], *Vincent Alcazer, MD PhD* [3,5], *Christophe Bouvier, DM* [6,3], *Julien Bollard, PhD* [7], *Audrey Lajoinie* [1], *Franck E. Nicolini, MDPhD* [7,3]

[1] RCTS, Lyon, France
[2] Novartis Pharma S.A.S, Rueil Malmaison, France
[3] Fi-LMC group, Lyon, France
[4] Laboratory of Hematology, Centre Hospitalier Lyon Sud, Pierre Benite, FRA
[5] Hematology department 1 G, Centre Hospitalier Lyon Sud, Pierre Benite, France
[6] Equipe Evaluation Médicale et Sarcomes (EMS), Centre Léon Bérard, Lyon, France
[7] Hematology Department, Centre Leon Berard, Lyon, France

Introduction
Over the past decade, the landscape of CML has been radically improved by the introduction of Tyrosine Kinase Inhibitors (TKI). However, since prolonged TKI treatment would reduce compliance, with associated risk of relapse and progression, and increased toxicities, patients (pts) who achieve and maintain a deep molecular response can now attempt TKI treatment discontinuation to achieve TFR. As TFR attempts may fail due to loss of major molecular response and imply restart of TKI therapy, it is essential to study individual factors associated with TFR outcome and to develop predictive models to provide TFR outcome likelihood with respect to pts' characteristics.

Methods
Predictive models of TFR failure at 6, 12, 24 and 36 months after TFR attempts were developed on 2 populations: (i) on CP- CML pts managed in a reference center for CML who attempted TFR between January 2003 and January 2023, named CML-pts and (ii) on those CML-pts for which an individual indirect linkage to nationwide claim database succeeded, named CML-linked-pts. The model was thus developed based (i) on clinical data exclusively (demographics, CML characteristics, TKI lines, molecular response), and (ii) on clinical data combined with claim data, which provided information on comorbidities, other treatments than TKI and any healthcare resource use (HCRU) of interest. Correlation and features' predictive power was assessed through Pearson's correlation and predictive power scores (PPS). Logistic regression model and machine learning algorithms (k-nearest neighbor, support vector machine, single decision tree, random forest, multilayer perceptron, and gradient boosting machine) were used. Nested k-fold cross-validation (CV) was used to assess the robustness of the algorithms over all data and hyperparameter optimization was performed. Features' contribution to predict TFR outcomes were assessed using SHAP ("SHapley Additive exPlanations") values.

Results
Analysis on (i) 208 CML-pts, who attempted 250 TFR - 208 first attempts (TFR $_1$), 42 second attempts (TFR $_2$). Median age was 59.6 (IQR:20.6) and 60.6 (IQR:18.2) years at TFR $_1$ and TFR $_2$, respectively, and M/F sex ratio was 1.17 and 1.1. The overall proportion of TFR failure was 35.6%, 43.6%, 46.8% and 48.8% at 6, 12, 24, and 36 months, respectively.

Analysis on (ii) 118 CML-linked-ptss, who attempted 134 TFRs (110 TFR $_1$ 24 TFR $_2$). Median age was 61.5 (IQR: 21.1) and 61.4 (IQR: 12.0) years at TFR $_1$ and TFR $_2$, respectively, and M/F sex ratio was 1.04 and 1.18. The proportion of TFR failure was 35.0%, 43.2%, 46.2% and 49.3% at 6, 12, 24 and 36 months, respectively.
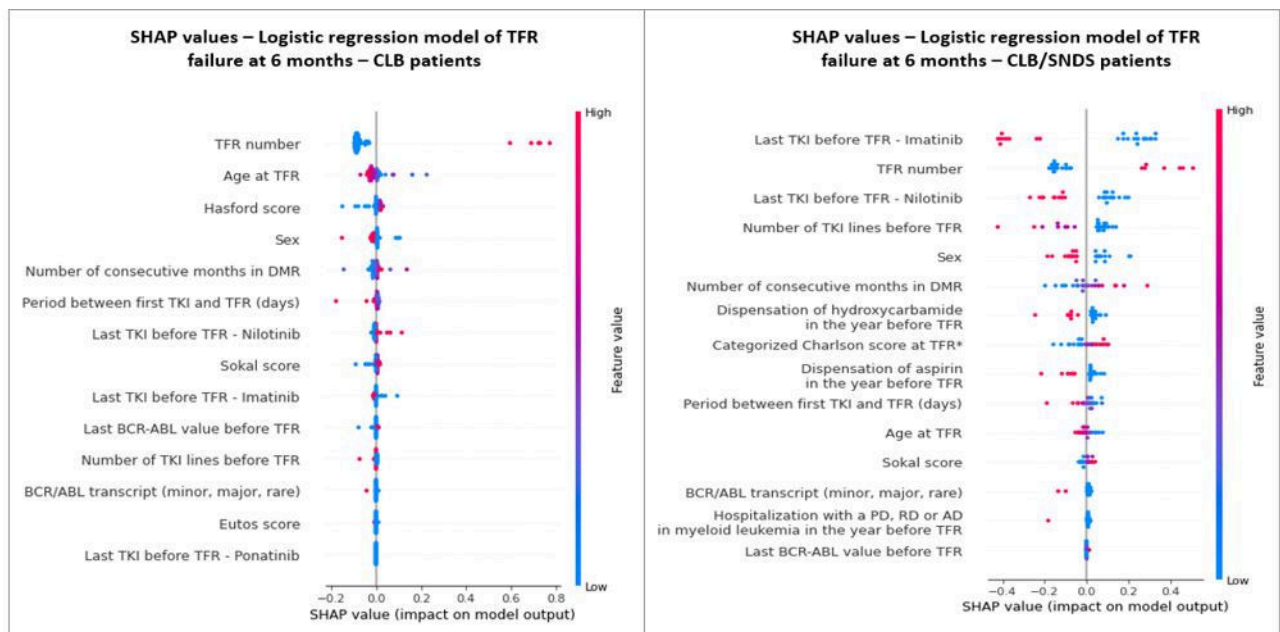
Studied predictive variables, in both populations (i) and (ii), had little correlation with TFR outcomes, ranging respectively from (i) -0.2 to 0.2 and (ii) -0.16 to +0.31, with poor predictive power (PPS from (i) 0 to 0.19, (ii) 0 to 0.25). Overall, predictive models hardly correctly predicted TFR failure - with sensitivity below 50% and were stronger at rather predicting absence of TFR failure - with specificity above 80%. Best performances were obtained with logistic regression model for both populations, for predicting at 6 and 12 months for population (i) and 6 months for population (ii). Accuracies were 72% for population (i) - with F1-score (i.e., model's ability to predict failure) of 50.0% at 6 months and 69.6% at 24 months -, and 77.8% for population

(ii) - with F1-score 66.7% at 6 months. SHAP values showed that attempting a second TFR was the most predictive factor of TFR failure for both populations (i) and (ii) (see Figure 1). Sensitivity analyses considering only TFR $_1$ in the models showed no improvement in performance.

Conclusion

Although this study was conducted on one of the largest TFR cohorts ever studied, prediction abilities of the models for TFR failure at the different time points were not sufficient to be proposed as a decision support tool, when developed with either clinical data or with both clinical and claim data. The models were indeed able to predict success but did not manage to find patterns to predict failure; only TFR number had predictive power of TFR outcome. Besides, claim data did not provide HCRU-related predictive factors. Among study limits, this monocentric cohort would reflect homogeneous practices in a same center. Analysis on larger cohorts of patients from different centers would potentially allow the capture of more variability in clinical practices.

*Categorized Charlson score is weighted by age
DMR: Deep Molecular Response (here, BCR-ABL ≤ 0.0032%)

Figure 1. SHAP values of logistic regression models of TFR failure at 6 months, for models developed on population (i) 208 CML-patients (left hand side), and on population (ii) 118 CML-linked-patients (right hand side)

Figure 1

https://doi.org/10.1182/blood-2023-179970