A machine learning usage to predict venetoclax efficacy and toxicity in chronic lymphocytic leukemia.

Aim of the project

The study aims at developing a prognostic and predictive model utilized for 1) identifying patients at risk for venetoclax therapy failure, and 2) identifying patients at high risk for developing treatment-specific adverse events. The major objective is to identify patients with a high probability of not achieving CR or CR with non-detectable MRD after completion of 12 courses of venetoclax-based therapy (VR or VenG). Secondary endpoints include prediction of the specific toxicities, most importantly neutropenia and neutropenic infections. The model should be based on an accessible in routine practice and easily defined set of data collected before therapy initiation and during the ramp-up period. Additionally, regarding the available data from the VERITA PALG-CLL5 clinical trial, the model will be enriched with a complete set of studied biomarkers including a wide panel of gene mutations assessed by NGS to verify if it may increase prediction efficacy. Ultimately, the created tool will be publicly available online as a www service (implementation in R Shiny and/or JavaScript).

Study description

We intend to analyze patient clinicopathological characteristics at treatment initiation and venetoclax ramp-up period in patients treated with venetoclax plus anti-CD20 antibody (rituximab or obinutuzumab) combination used for treatment of newly diagnosed CLL. The data that we will analyze have previously been collected as part of two clinical studies conducted by the Polish Adult Leukemia Group (PALG), one prospective non-commercial controlled clinical trial and one retrospective analysis. PALG, as the data owner, has consented to the use of mentioned data for this project. For developing the machine learning algorithms/models we will use the data taken from 103 previously untreated patients with a diagnosis of CLL /small lymphocytic lymphoma (SLL) who had an indication to start therapy according to published 2018 IWCLL NCI-WG Criteria and were enrolled to the VERITA PALG-CLL5 study, to which recruitment has just ended. For verification of the algorithm in real-world patients (outside clinical trials) we will use the data obtained within the PALG retrospective study "Analysis of the effectiveness and toxicity of treatment according to the venetoclax and Obinutuzumab (VenG) regimen in the first line of chronic lymphocytic leukemia treatment in clinical practice in Poland". Both models are to be developed. First, will be based solely on patients' clinical data and routine cytogenetic risk profile; second will incorporate the mutational status of selected genes obtained via next-generation sequencing. Both initial models will be verified in real-world CLL patients treated with venetoclax-based treatment outside clinical trial settings.

Why is this project important?

The management of patients with CLL continues to improve with novel therapies and multiple biological and genetic variables. Although significant work on their validation is still required before they can be used in a routine clinical setting, they may improve the accuracy of current prognostic models. The natural history of CLL varies, with a wide spectrum ranging from a slow, indolent course to rapid disease progression. Several decades ago, staging systems were developed incorporating simple clinical parameters such as complete blood counts and findings from physical examinations. These systems are widely accepted as validated prognostic tools, but limitations were identified on their usefulness when dealing with individual patients. There is a need to develop easy to use, practical prognostic models that would predict efficacy and complications of new standard-of-care targeted therapies for CLL, such a venetoclax, with higher accuracy than existing prognostic parameters and scores that were developed for chemoimmunotherapy. It is noteworthy that the use of machine learning algorithms in the prediction of treatment response in CLL has appeared in scientific papers only in recent years. It became standard to analyze typical CLL diagnostic parameters, and the models created were based on correlations. Therefore, our approach to developing such an algorithm for venetoclax plus anti-CD20 regimens for first-line CLL therapy seems rational and feasible.

Expected results

We hypothesize that using machine learning methods to analyze molecular, laboratory and clinical data available before therapy initiation and laboratory results obtained during the early phase of treatment with venetoclax (routine dose escalation phase called the ramp-up period) will allow us to develop a web-based tool for predicting the efficacy and toxicity of venetoclax-based therapy.