Predictive Modeling as an OHDSI Network Study: Treatment Non-Response in NSCLC patients treated with ALK Inhibitors

Aaron Galazky1, Greg Kiebaro2, Eldar Alakhyverdiev3, Yuri Khoma4 and Christian Rein3
1Takeda Pharmaceuticals, Cambridge MA USA; 2Odysseus Data Services, Cambridge MA USA; 3QuintilesIMS, Cambridge MA USA

Abstract
In this study, we determine the rate of non-response to treatment with ALK inhibitors in patients of Small Cell Lung Cancer (NSCLC). We also develop a predictive algorithm for identifying patients at high risk of treatment non-response to these agents. Both investigations are conducted via OHDSI network study, i.e. the R developed on a local database is distributed for execution against a suite of US database assets remotely using the ARACHNE study execution tool.

Introduction
Targeted therapy is now standard of care (SOC) for a subset of patients whose tumors harbor actionable abnormalities, exemplified by Non-Small Cell Lung Cancer (NSCLC) mutations. EGFR mutations are the most common of these mutations and are effectively treated using TKI inhibitors, with patients achieving significant response rates. However, ALK mutations are less common, occurring in about 5% of NSCLC tumors. Tumors with ALK mutations have significantly lower response rates to TKI therapy. As an area of unmet need, Takeda is interested in developing novel therapies targeted to reduce non-NSCLC patients with less common mutations. To support these efforts, Takeda is interested in characterizing NSCLC patients, the risk ratio rate of treatment non-response to ALK inhibitors, which is increasingly being used in NSCLC, as well. This information is used to develop a predictive modeling algorithm for identifying patients at high risk of treatment non-response to these agents. The risk of the algorithm would be used to identify patients likely to benefit from novel therapeutics through clinical parameters, thereby supplementing genetic/biomarker based approaches. Such an approach has potential applications to characterizing population needs, clinical study recruitment, and treatment decision-making for patients with NSCLC.

Study overview
This study developed a risk model and used this model to predict treatment outcome for patients with locally advanced and metastatic NSCLC, constituting a retrospective, observational, non-control cohort study. We used OHDSI PatientLevelMatcher package to develop the risk model and to predict treatment outcome for patients with locally advanced and metastatic NSCLC, constituting a retrospective, observational, non-control cohort study.

Study Design
All subjects in the database were included into the model following criteria (Fig 1): Inclusion criteria:
- Exposure to at least 1 ALK inhibitors
- Underwent chemotherapy or radiotherapy
- At least 182 days of observation time prior to the index date
- At least 10 years of age at the first diagnosis date (diagnosis is locally advanced or metastatic NSCLC)
- Absence of second malignancies during the washout period
- At least 2 month gap in treatment before the previous regimen was allowed prior to the index date.

Outcome definition
The primary outcome for each treated patient is response or non-response to treatment. Patients were assumed to be a non-responder if:
- A new line of therapy started after the washout gap
- Treatment with another standard chemotherapy was violated
- Surgical or radiologic procedures were administrated
- If none of these occurred within a window of observation the patient was considered a responder. As the data set had privacy initiatives on de-identifying information, initial attempts could not be confirmed.
- Response was observed in three different observation windows (Fig 2.): at any time after index date, observation window 1, 6 months after the index date (observation window 2), and 1 year after the index date (observation window 3).

For feasibility and development of the study costs, we used a commercial US claims database.

As a next step, we are planning to post the OHDSI call for collaborators to participate in a public network study.

Conducting Distributed Network Study
We are planning to execute the study and demonstrate results for a number of US data assets derived from payer-based claims, provider-based claims and EHR systems.

Results
The following shows the preliminary results generated from the databases used for development:
473 patients could be identified in the development database. In which were distributed into responders and non-responders as follows:
- Observation window 1: 197 responders and 276 non-responders
- Observation window 2: 298 responders and 175 non-responders
- Observation window 3: 232 responders and 241 non-responders

Fig 2. Observation window definition and patient distribution within them

Model used:
Naive Bayes, Gradient Boosting machine, Random Forest and neural network were trained for each observation window. The preliminary results for the predictive model were as follows:

Conclusion
It is possible to predict treatment outcome for NSCLC patients, but the sample size was rather small to each conclusion. In particular, the fact that different methods generated nearly the same accuracy (Fig 3.) can be attributed to them probably generating the same effective model. The future work is expected using the larger OHDSI network. We also intend to expand our modeling efforts to include other targeted therapies, such as EGFR inhibitors.