DATAVANT

Datavant Life Sciences Case Study:

Using De-identified Linked Real-World Data to Create an External Control Arm in Non-Small Cell Lung Cancer

Client Situation

A Pharmaceutical company designed its Phase II study of a novel agent to treat metastatic non-small cell lung cancer driven by a rare subtype of EGFR genetic mutation. This type of lung cancer only affected a small subset of EGFR patients, therefore designing studies and finding patients for clinical trials was difficult. The client decided that constructing a control arm out of real-world electronic health record data would be the best approach to the trial.

Client Need

For the external control arm, the company needed a cohort of patients with the specific subtype of metastatic EGFR non-small cell lung cancer whose disease progressed despite taking platinum-based chemotherapy. The client found that to create a large enough cohort of patients meeting the clinical criteria, they needed to aggregate electronic health record data (EHR) from three different commercial real world data vendors. When they approached the FDA with the study proposal, FDA made their acceptance of the aggregated data contingent upon the company demonstrating that the data across the three vendors represented unique patients (i.e., no duplicate patients across data sets).

How Datavant Helped

Each one of the Oncology EHR vendors tokenized their data using the Datavant platform and delivered the de-identified data to the company. The company was able to compare de-identified tokens across the three data sets and ensure the aggregated data set only included unique patients. The ability to de-duplicate these three data sets and ensure they had a cohort of unique patients enabled the company to move forwards with the Phase II pivotal trial.