Lightning Talks
1:00-2:00pm

Towards an OHDSI-based Data-Sharing Infrastructure in Germany Orchestration
The German Federal Ministry of Education and Research recently started a large data integration and data sharing research initiative, in order to improve the reuse of data from patient care and translational research. MIRACUM (Medical Informatics for Research and Care in University Medicine) is the largest of four funded consortia and brings together more than 18 institutions within 5 states of Germany. It will use the OMOP common data model and the OHDSI tools for harmonizing the terminologies used as well as facilitating the federation of research analyses across institutions. The talk will introduce the national initiative and the MIRACUM project, highlight results of an early, federated analysis using OMOP and name future work and challenges for use of OHDSI in Germany.
Presenter: Dr. Martin Sedlmayr, Department of Medical Informatics, Friedrich Alexander-University Erlangen-Nürnberg

Transforming Electronic Health Records from Swedish registers to the OMOP CDM
Sweden, together with other Nordic countries, has one of the best centralized Electronic Health Records (EHR) in Europe. At The Hyve we created an Extract Transform Load procedure to transform Swedish inpatient, outpatient, drug, death and population registers into the OMOP Common Data Model. This ETL procedure was executed remotely for an Atrial Fibrillation study with about 468k patients. During this talk, we will highlight some translation challenges we encountered, both for incompatible data structures and vocabulary mappings of Swedish coding systems.
Presenter: Maxim Moinat, The Hyve

The feasibility of utilizing the OHDSI network to generate large-scale evidence of the safety of biologics
We have proposed a large-scale post-market surveillance program for biologics for funding from the Australian National Health and Medical Research Council (NHMRC). The project will be a collaboration between OHDSI and the Asian Pharmacoepidemiology Network (AsPEN) to describe the patterns of use of biologics, the safety of biologics and the identification of factors that place patients at increased risk of these events. In response to reviewers concerns of the sample size required for such studies, given the often rare adverse events, a feasibility assessment was performed across the OHDSI community. This presentation will describe the process of performing a rapid feasibility assessment for a study to generate evidence of the safety of biologics and summarize the results in the context of current small scale observational studies in this area.
Presenter: Nicole Pratt, University of South Australia

High-throughput phenotyping using imperfectly-labeled training data

The widespread adoption of electronic medical records provides a unique opportunity to accelerate research. To use these data for research, it is necessary to identify patients with phenotypes of interest and construct cohorts for observational studies. Traditional methods of phenotyping involve extensive rule-based definitions, which require a considerable amount of time to create. Recently, supervised learning techniques have been used to build phenotype definitions in the form of classifiers. The **A**utomated **PH**enotype **R**outine for **O**bservational **D**efinition, **I**dentification, **T**raining and **E**valuation (**APHRODITE**) is an R-package phenotyping framework that uses imperfectly labeled data to train classifiers, and has been shown to achieve good classification performance for type 2 diabetes mellitus and myocardial infarction. In previous work, the labeling heuristic used in APHRODITE have been based on mentions of phrases in textual data. In this study, we trained high-precision classifiers for four distinct phenotypes (type 2 diabetes mellitus, glaucoma, epileptic seizure, and peripheral vascular disease) using a labeling function based on multiple mentions of disease-specific codes. APHRODITE classifiers constructed with such imperfectly labeled data show significant improvements in recall compared to multiple-mention based labeling functions. These results demonstrate that it is possible to build phenotype classifiers in a high-throughput manner using the APHRODITE framework even in the absence of textual data to drive the labeling of the training set.
Presenter: Mehr Kashyap, Stanford University

Uncovering Exposures Responsible for Birth Season – Disease Effects: A Global Study

Birth month and climate impact lifetime disease risk while the underlying exposures remain largely elusive. We seek to uncover distal risk factors underlying these relationships by probing the relationship between global exposure variance and disease risk variance by birth season. This study utilizes Electronic Health Record data from six sites representing ten and a half million individuals and three countries (USA, South Korea and Taiwan). Three of these sites are OHDSI collaborators. We obtained birth month – disease risk curves from each site in a case-control manner. Next, we correlated each birth month – disease risk curve with each exposure. A meta-analysis was then performed of correlations across sites. This allowed us to identify the most significant birth month – exposure relationships supported by all six sites while adjusting for multiplicity. We also successfully distinguish relative age effects (a cultural effect) from environmental exposures. Attention Deficit Hyperactivity Disorder was the only identified relative age association. Our methods identified several culprit exposures that correspond well with the literature in the field. We will discuss these findings at our talk.
Presenter: Mary Regina Boland, University of Pennsylvania

Learning Effective Clinical Treatment Pathways for Type-2 Diabetes from Observational Data

The treatment guidelines for Type-2 Diabetes (T2D) remain ambiguous for second line therapy. When the initial prescription of Metformin results in adverse effects or patient-level hemoglobin A1c goals are not met, there is little guidance on which second line drug to initiate. Considerable variation has been reported in the practice of managing T2D across healthcare systems globally. We developed an approach to learn effective clinical treatment pathways in

T2D from observational clinical data. Data from over 101 million patients in seven healthcare systems across four countries was organized into the OMOP common data model. Treatment pathways, representing T2D medication trajectory of each patient, were built and compared across healthcare systems to understand variation in treatment response. The drug efficacy of second-line drugs among Sulfonylureas, DPP4-Inhibitors and Thiazolidinediones was assessed for reduction in HbA1c <= 7%, occurrence of myocardial infarction, as well as kidney and eye related disorders. The study was executed and assessed for reproducibility at each of the healthcare system individually. The results obtained from each healthcare system were combined via a meta-analysis. DPP4-Inhibitors were identified as best second-line drug after Metformin in reducing the HbA1c levels of patients with T2D. The second line drugs examined did not differ in their effect in preventing myocardial infarction, kidney or eye related disorders.
Presenter: Rohit Vashisht, Stanford University

Comparison of combination treatment in hypertension
Currently, the American and European guidelines recommend initial combination treatment in high-risk hypertension patients. However, the real-world evidence of head-to-head comparison among the recommended regimens is still limited. We aim to compare the therapeutic effectiveness of combination regimens between patients initiating dual antihypertensive treatment. From Korean National Health Insurance sample cohort database, we identified eligible patients without previous history of cardiovascular disease who were started on and received prescription of dual anti-hypertensive treatment for more than 180 days between 2003 and 2012. The patients were matched for each comparison set by large scale propensity score matching. Primary end point was all-cause mortality. The results suggest that there is no significant difference of all-cause mortality among recommended dual combination treatment regimen.
Presenter: Seng Chan You, Ajou University