Personalizing Background Risk Estimates for Outcomes of Interest associated with Covid-19 Vaccination

Jenna Reps, Patrick Ryan
Janssen R&D Titusville, NJ, USA

Background

Covid-19 vaccinations are being administered at scale around the world. Numerous outcomes have been identified by the FDA as outcomes of interest to monitor. At present the best personalized risk estimates for the outcomes are based on population-based averages. However, it may be possible to develop patient-level prediction models that can personalize risk estimates based on a patient's demographics and medical history. It is currently unknown i) whether the covid-19 vaccination outcomes of interest can be predicted using observational data, ii) whether certain datasets may result in better performing models and iii) whether these models generalized across target populations and outcome phenotype definitions.

Methods

In this paper we use the OHDSI PatientLevelPrediction [1] framework to implement a large-scale prediction analysis where we investigate three target populations (random visit, Jan 1st set date and influenza visit), nine outcomes (hemorrhagic stroke, non-hemorrhagic stroke, myocardial infarction, anaphylaxis, appendicitis, disseminated intravascular coagulation, encephalomyelitis, Guillain Barre syndrome and pulmonary embolism) with 3-4 phenotypes per outcomes, seven databases (Optum EHR, Optum Claims, CCAE, MDCD, MCDR, JMDC, IQVIA German) and two covariate settings (age/sex only and age/sex/medical history in past 365 days). We develop models for all possible combinations of target population, outcome phenotype, database and covariate setting, in total this corresponds to > 1000 models. No patients with COVID-19 vaccinations were included in this study, but the study will help inform whether the COVID-19 vaccination outcomes of interest are generally predictable using observational healthcare data.

Model development followed a cohort design where we predict each outcome phenotype occurring within 1-year of index. We include patients that are right censored during the 1-year follow-up. We randomly sampled 2 million patients for each target population with more than 2 million patients. LASSO logistic regression models were developed using an 80% train and 20% test split with 3-fold cross-validation on the train data to select the optimal regularization parameter. The internal validation in terms of area under the receiver operating characteristics (AUROC) discrimination and calibration-in-the-large (mean predicted risk for the population vs observed risk) were calculated using the test set.

Model generalizability was calculated by externally validating the models across all possible combinations of target population, outcome phenotype and database. This corresponds to > 60,000 model validations. The external validation also investigated AUROC discrimination and calibration-in-the-large (mean predicted population risk vs observed population risk). External validation at scale is made possible due to the OHDSI standardizations [2].

Results

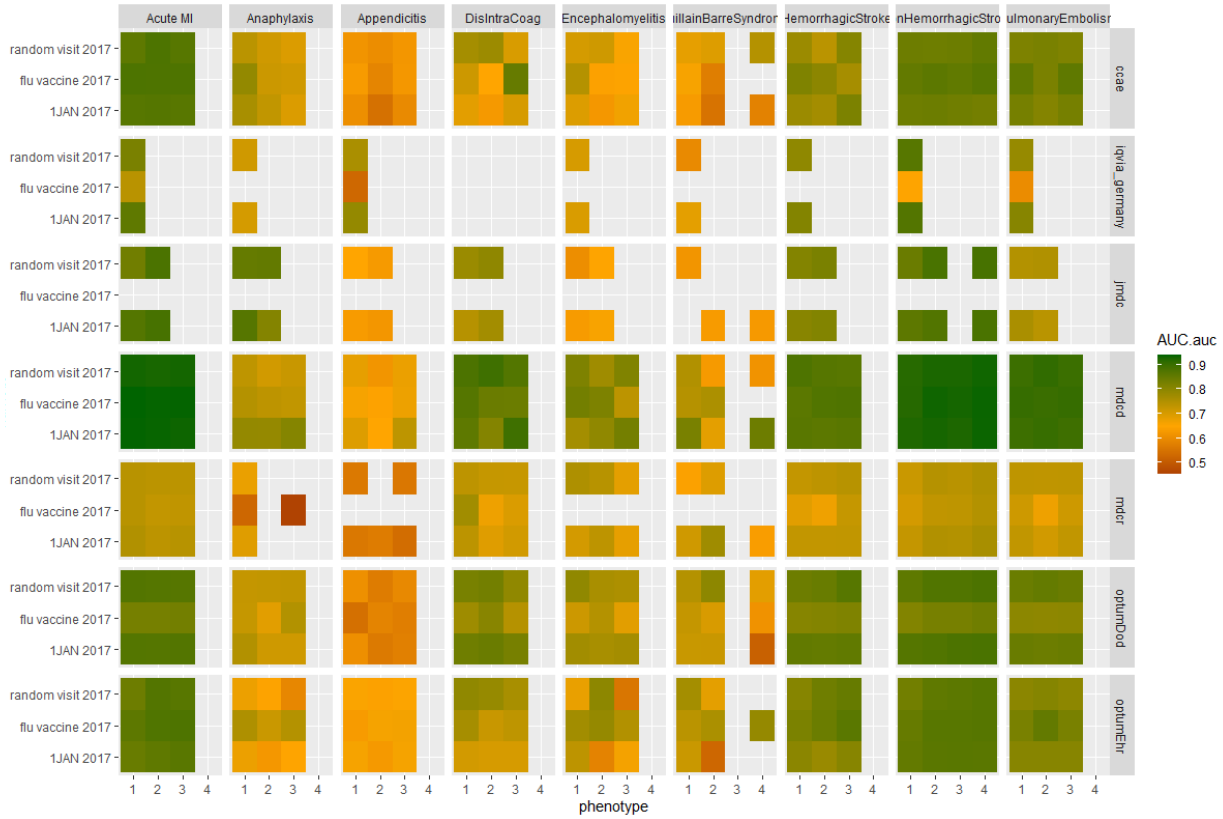

Figure 1- Heat map showing the internal AUROC values across target populations (y-axis), phenotypes (x-axis) group by outcomes (x group) and databases (y group) when developing models using all demographics/conditions/drugs in the prior year.

Figure 1 shows the internal discriminative performances across databases, target cohorts, outcomes and outcome phenotypes when using demographic plus condition/drug covariates. Acute MI, hemorrhagic and non-hemorrhagic stroke and pulmonary embolism appear to be highly predictable across databases, target cohorts and outcome phenotypes. Some outcomes were more predictable in certain databases, such as disseminated intravascular coagulation, which was more predictable in MDCD and Optum Claims. The predictability of encephalomyelitis and Guillain Barre syndrome appear to depend on the outcome phenotype and target databases. In the IQVIA Germany data it seems the target population impacted predictability, where it was harder to predict outcomes in patients given the flu vaccination. This may be due to older/sicker people getting the flu vaccination in Europe. In Optum EHR some outcomes, such as anaphylaxis and encephalomyelitis, appear easier to predict after the flu vaccination. Some phenotypes were not suitable for certain databases, for example, in JMDC the flu vaccination target cohort did not identify any patients.

Figure 2 illustrates the difference in performance of models developed using medical history and age/sex minus the models that only included age/sex. Figure 2 shows that certain outcomes, such as anaphylaxis, encephalomyelitis and Guillain Barre syndrome are more predictable when conditions/drugs are included as covariates. The cardiovascular outcomes appear predictable just using

age/sex, except in MDCR where medical history helped discriminate people likely to experience the outcomes. Some models performed better when age/sex were the only covariates (e.g., target 1st Jan 2017, outcome disseminated intravascular coagulation in CCAE). These were often target and outcome pairs that were difficult to predict (orange in Figure 1).

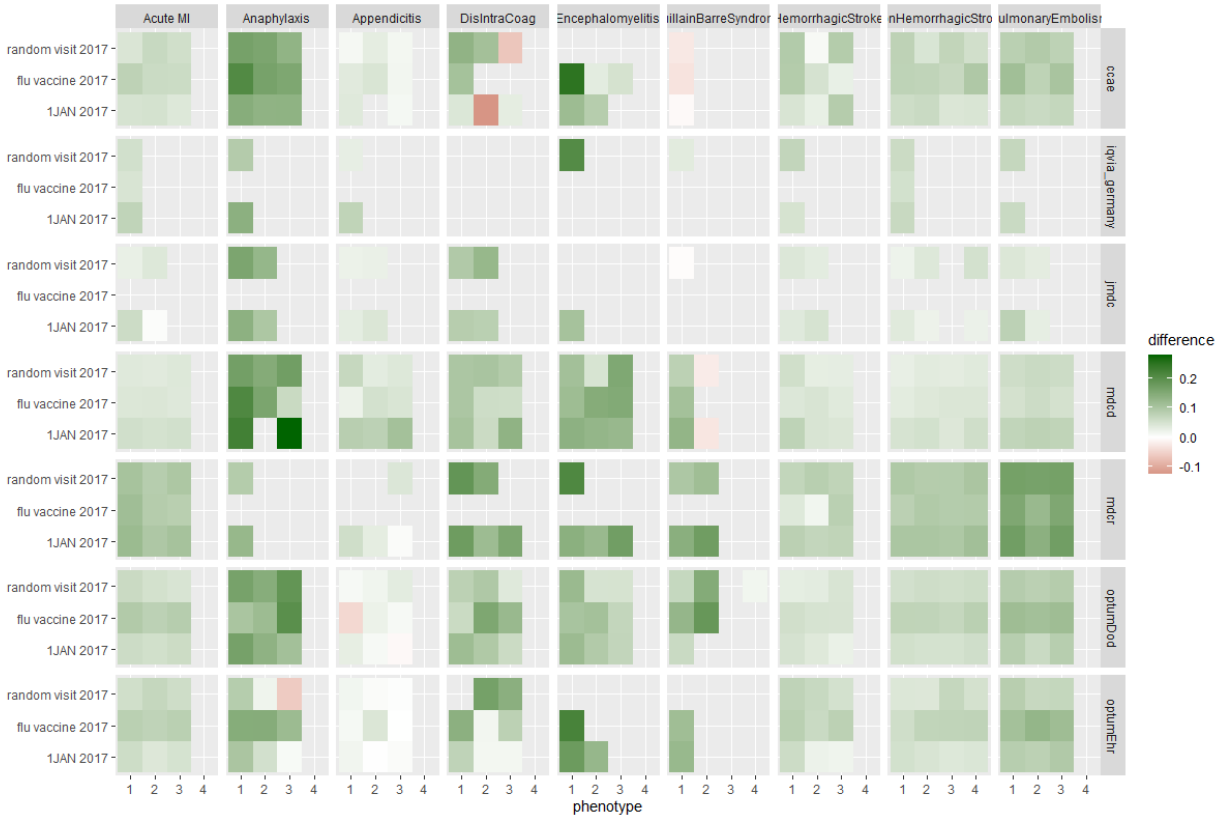

Figure 2- Heat map showing the difference in internal AUROC values for the medical history/age/sex models minus the age/sex only models across target populations (y-axis), phenotypes (x-axis) group by outcomes (x group) and databases (y group).

Conclusion

The results show that most of the FDA COVID-19 vaccine outcomes of interest can be predicted using claims and EHR data. Some of the outcomes are highly predictable with just age whereas others require medical history covariates. Many models were stable across target cohorts and outcome phenotypes, suggesting that may be likely to transport into a clinical setting. These models could be used to personalize the risk that a patient experiences one of these outcomes after the COVID-19 vaccination. This could be used to identify whether an outcome occurred unexpectedly. Prediction models do not provide any causal inference information and should not be used to make causal statements.

References/Citations

1. Reps, J.M., Schuemie, M.J., Suchard, M.A., Ryan, P.B. and Rijnbeek, P.R., 2018. Design and implementation of a standardized framework to generate and evaluate patient-level prediction models using observational healthcare data. *Journal of the American Medical Informatics Association*, *25*(8), pp.969-975.
2. Reps, J.M., Williams, R.D., You, S.C., Falconer, T., Minty, E., Callahan, A., Ryan, P.B., Park, R.W., Lim, H.S. and Rijnbeek, P., 2020. Feasibility and evaluation of a large-scale external validation approach for patient-level prediction in an international data network: validation of models predicting stroke in female patients newly diagnosed with atrial fibrillation. *BMC medical research methodology*, *20*, pp.1-10.