

Making patient-level predictive network study packages

Jenna Reys, Martijn J. Schuemie, Patrick B. Ryan, Peter R. Rijnbeek

2021-12-16

Contents

1	Introduction	2
2	Useful publication	2
3	Main steps for running a network study	2
3.1	Step 1 – developing the study	2
3.2	Step 2 – implementing the study part 1	2
3.3	Step 3 – implementing the study part 2 (make sure the package is functioning as planned and the definitions are valid across sites)	2
3.4	Step 4 – Publication	2
4	Package Skeleton - File Structure	2

1 Introduction

The OHDSI Patient Level Prediction (PLP) package provides the framework to implement prediction models at scale. This can range from developing a large number of models across sites (methodology and study design insight) to extensive external validation of existing models in the OHDSI PLP framework (model insight). This vignette describes how you can use the `PatientLevelPrediction` package to create a network study package.

2 Useful publication

The open access publication A standardized analytics pipeline for reliable and rapid development and validation of prediction models using observational health data details the process used to develop and validate prediction models using the OHDSI prediction framework and tools. This publication describes each of the steps and then demonstrates these by focusing on predicting death in those who have covid-19.

3 Main steps for running a network study

3.1 Step 1 – developing the study

- Design the study: target/outcome cohort logic, concept sets for medical definitions, settings for developing new model or validation of adding existing models to framework. Suggestion: look in literature for validated definitions.
- Write a protocol that motivates the study and provides full details (sufficient for people to replicate the study in the future).
- Write an R package for implementing the study across diverse computational environments [see guidance below for structure of package and use the skeleton github package here: <https://github.com/OHDSI/SkeletonPredictionStudy>]

3.2 Step 2 – implementing the study part 1

- Get contributors to install the package and dependencies. Ensure the package is installed correctly for each contributor by asking them to run the `checkInstall` functions (as specified in the `InstallationGuide`).
- Get contributors to run the `createCohort` function to inspect the target/outcome definitions. If the definitions are not suitable for a site, go back to step 1 and revise the cohort definitions.

3.3 Step 3 – implementing the study part 2 (make sure the package is functioning as planned and the definitions are valid across sites)

- Get contributors to run the `main.R` with the settings configured to their environment
- Get the contributors to submit the results

3.4 Step 4 – Publication

The study creator has the first option to be first author, if he/she does not wish to be first author then he/she can pick the most suitable person from the contributors. All contributors will be listed as authors on the paper. The last author will be the person who lead/managed the study, if this was the first author then the first author can pick the most suitable last author. All authors between the first and last author will be alphabetical by last name.

4 Package Skeleton - File Structure

- `DESCRIPTION`: This file describes the R package and the dependencies
- `NAMESPACE`: This file is created automatically by Roxygen

- Readme.md: This file should provide the step by step guidance on implementing the package
- R
- helpers.r: all the custom functions used by the package should be in this file (e.g., checkInstall)
- main.r: this file will call the functions in helpers.r to execute the full study
- submit.r: this file will be called at the end to submit the compressed folder to the study creator/manager.
- Man: this folder will contain the documentation for the functions in helpers.r (this should be automatically generated by roxygen)
- Inst
- sql/sql_sever * targetCohort: the target cohort parameterised sql code * outcomeCohort: the outcome cohort parameterised sql code
- plp_models: place any PLP models here
- Extras