PlamodeSim: A package to simulate data with a known outcome generating process

**Background:** In order to compare new or existing methods we aimed at developing a package that simulates data with a known outcome generating process while staying close to the real life data. To this extent, we develop the R package *PlamodeSim*.

**Result 1:** Code for simulating new outcomes from an existing and a modified prediction model.

**Result 2:** Distribution of baseline risk for a binary outcome in the original (Eunomia) and simulated datasets.

```r
# Derived model
plpPrediction <- PatientLevelPrediction::predictPlp(
  plpModel = plpResult$model,
  plpData = plpData,
  population = population)
newOutcomesFittedModel <- PlasmodeSim::newOutcomes(
  noPersons = 2000,
  probs = plpPrediction)

# Modified model
unfittedModel <- plpResult$model$coefficients
unfittedModel[3:5, 1] <- 0.4
unfittedModel[1, 1] <- -0.4
plpModelunfitted <- PlasmodeSim::makeLogisticModel(unfittedModel)
newprobs <- PatientLevelPrediction::predictPlp(
  plpModel = plpModelunfitted,
  plpData = plpData,
  population = population)
newOutcomesUnfittedModel <- PlasmodeSim::newOutcomes(
  noPersons = 2000,
  probs = newprobs)
```

**Methods**

The most important features implemented are:
- Functions to simulate datasets by sampling the original data with replacement.
- Functions that simulate binary or survival outcomes.
- Functions for visual diagnostics.

Gidius van de Kamp, Alexandros Rekkas