SPONSOR	Roland MARION-GALLOIS
SUPERVISOR	Barbara ROSETTANI
DEPARTMENT	Biometrics and Data Science, MAMA & Pediatrics, Boudry
PREREQUISITES	MSc or Ph.D. student in (Bio-) Statistics. Experienced in SAS and/or R and/or Python

Description of the project and work required (including references)

The analyses of adverse events (AEs) plays an important role in the regulatory process of drug approval, as well as in the benefit assessment of drugs after approval. In case of varying follow-up times, methods based on survival time analyses are preferred compared with analyses based on contingency tables. In this setting, competing risks models are recommended for the estimation of the cumulative incidence functions[1,2].

Methods for the estimation of the cumulative incidence function and for studying the covariate effects on the cumulative incidence function are available. One of the standard approaches is to model cause-specific hazards for all causes using the Cox proportional hazards model (or other semiparametric models, or parametric models e.g Weibull, lognormal models). Another method is the Fine-Gray model that directly assesses the effect of covariates on the sub-distribution function. Confusion still exists as to how choose amongst available methods and how to interpret effect estimates. A new approach to competing risks is using random forests that is a fully non-parametric method and can be used for selecting event-specific variables and for estimating the cumulative incidence function.

However, those methods have not used a formal framework for characterizing causal effects. In this proposed project the aim is to explore methods for accounting for potential confounding in competing risk settings, in the causal framework, the parametric g-formula and IPW estimators will be used to estimate the risk of the competing event and the corresponding total treatment effects.

To make the approach robust to model-misspecification, we will attempt to use Superlearner, an ensemble of methods.

The above methods will be applied to study PASS-001, an observational post authorization safety study with primary endpoint characterize and determine the incidence of adverse events of special interest. The study includes 2150 subjects in the Lenalidomide cohort and 1479 in the Background cohort. It would be of great value to make an exhaustive analysis on the most important adverse events using competing risk methodology.

Skills/Knowledge Required (technical and soft skills)

Master or PhD student in Statistics (ideally Bio-Statistics) with good understanding of basic methodology of clinical research. Excellent programming skills and autonomy in using R and/or SAS.

Team player. Good organization and communication skill.

Anticipated/expected outcome of the project (how would you define a successful internship?)

The student will have to deliver an internship report, in alignment with requirements from his/her university. For BMS internal use, the student will have to deliver a guidance document and PowerPoint presentation.

Benefits to the candidate

Internship should be considered part of the student education.
The intern will gain working experience of applying causal inference and machine learning in the context of pharmaceutical industry.
Internship outcomes may lead to publications and further researches.

Other relevant items

Offer opened only to University/School students as part of their studies.
Internship is tripartite agreement/commitment between BMS, the student and the university.

Company: Celgene, a Bristol Myers Squibb company

Location: Boudry, Switzerland.

Internship period: According to the School/University rules.

Duration: 3 to 6 months minimum. According to School/University rules.