PRecIsion Medicine in CardiomyopathY (PRIMaCY)

Statistical Analysis Plan

July 21, 2020

All statistical analyses were performed using R version 3.4.1 with survival, multiple imputation by chained equations, mboost, riskRegression, and mlr packages. Clinical characteristics were summarized using descriptive statistics. Continuous variables were summarized using median and interquartile range, whereas categorical variables were summarized using frequencies and proportions. The primary outcome was time to an SCD event during 5-year follow-up defined as a composite outcome of SCD, resuscitated sudden cardiac arrest (including sustained VT/ventricular fibrillation), and aborted SCD events, that is, appropriate shock in a patient with a primary prevention ICD. Event-free survival was estimated by using the Kaplan-Meier method as was the cumulative proportion of SCD events over time.

Model Development and Performance Assessment

General Modeling Approach

To generate a risk prediction model, the relative contribution of the clinical and genetic predictors was quantified by using a competing risk model for SCD events with non-SCD death as the competing risk. This analysis was implemented by using cause-specific regression. To ensure that the quantification was pertinent to the clinical context, patients were administratively censored at 5 years from first evaluation. Two models were considered: a clinical-only model and a clinical/genetic model that included genotype status of the patient. Patients harboring a pathogenic/likely pathogenic causal variant on genetic testing, that is, genotype-positive patients, were compared with patients without a causal variant on testing. We applied a model-boosting algorithm to the cause-specific hazard regression to identify and quantify the association between the candidate risk factors and the composite outcome. For continuous risk factors, the associations were modeled without the assumption of linearity using penalized b-splines. The estimated nonlinear associations were summarized graphically. On the fitting of the 2 boosted models, we then obtained the linear predictors and baseline hazard for the SCD events.

Model Tuning

To tune the boosting model, we applied 10 repeated 4-fold cross-validations to determine the hyperparameter of the boosting model. In this exercise, we considered a small learning rate (step size) of 0.01 and a large number of steps. Within each cross-validation iteration, we imputed the missing values by using multiple imputation by chained equations before training the model. The plausibility of missing at random assumed in the imputation depends on the candidate risk factors. As long as the candidate risk factors considered in the study are comprehensive, as in our study, this assumption is reasonable. Next, we applied the trained model and calculated the c-statistic on the cross-validation data for each step. The optimal number of steps (ie, the value of the boosting model hyperparameter) was the one that maximized the c-statistic averaged over all cross-validation iterations. The final model was then fitted using all data with the optimal hyperparameter value estimated in the model-tuning exercise. Given the competing risk data, separate cause-specific hazard regression models were fitted and tuned: 1 for SCD events and the other for non-SCD deaths. Because of a small number of patients with non-SCD deaths, we only applied 3-fold cross-validation to identify and quantify the risk factors associated with non-SCD deaths.

Model Performance

On model fitting, both prognostic index and the 5-year cumulative proportion of SCD events were quantified for each patient. The discriminatory power of the final model was quantified using the c-

statistic for competing risk models. Model calibration was assessed by stratifying patients into 3 risk groups based on tertiles of the predicted 5-year probability of a SCD event and by creating a calibration curve to show the relationship between the observed and predicted 5-year probability of events.

Model Validation

The final model was externally validated using an independent replication cohort. The Harrell c-statistic was calculated, and calibration curves were constructed using quantiles of predicted risk as described above.