Leveraging Real-World Data for EMA Qualification of a Model-Based Biomarker Tool to Optimize Type-1 Diabetes Prevention Studies

Jagdeep T. Podichetty[1,*], Patrick Lang[1], Inish M. O'Doherty[1], Sarah E. David[1], Rhoda N. Muse[1], Stephen R. Karpen[1], Laura Sue Song[1], Klaus Romero[1] and Jackson K. Burton[1] on behalf of the Type-1 Diabetes Consortium (T1DC)

The development of therapies to prevent or delay the onset of type 1 diabetes (T1D) remains challenging, and there is a lack of qualified biomarkers to identify individuals at risk of developing T1D or to quantify the time-varying risk of conversion to a diagnosis of T1D. To address this drug development need, the T1D Consortium (i) acquired, remapped, integrated, and curated existing patient-level data from relevant observational studies, and (ii) used a model-based approach to evaluate the utility of islet autoantibodies (AAs) against insulin/proinsulin autoantibody, GAD65, IA-2, and ZnT8 as biomarkers to enrich subjects for T1D prevention. The aggregated dataset was used to construct an accelerated failure time model for predicting T1D diagnosis. The model quantifies presence of islet AA permutations as statistically significant predictors of the time-varying probability of conversion to a diagnosis of T1D. Additional sources of variability that greatly improved the accuracy of quantifying the time-varying probability of conversion to a T1D diagnosis included baseline age, sex, blood glucose measurements from the 120-minute timepoints of oral glucose tolerance tests, and hemoglobin A1c. The developed models represented the underlying evidence to qualify islet AAs as enrichment biomarkers through the qualification of novel methodologies for drug development pathway at the European Medicines Agency (EMA). Additionally, the models are intended as the foundation of a fully functioning end-user tool that will allow sponsors to optimize enrichment criteria for clinical trials in T1D prevention studies.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ Currently, development of therapies to prevent or delay the onset of type 1 diabetes (T1D) remains challenging, and there is a lack of qualified biomarkers for patient selection and quantification of risk of conversion to stage 3 T1D.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ This work leveraged existing patient-level data from three T1D observational studies to develop a time-to-event model for predicting the time-varying probability of T1D diagnosis, based on islet autoantibody (AA) seropositivity and other patient features, for the purposes of optimizing clinical trial design in T1D prevention studies.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ The developed models demonstrates that islet AAs are statistically significant predictors of the time-varying probability of conversion to a diagnosis of T1D during a reasonable duration for a T1D prevention trial, representing adequate underlying evidence for their use as enrichment tools in clinical trials.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

☑ Following the European Medicines Agency (EMA) qualification of the islet AAs, the model will be made publicly available to sponsors and regulators to help expedite drug development to prevent T1D.

Type 1 diabetes (T1D) is a chronic autoimmune disease caused by progressive destruction of insulin producing β-cells in the islets of Langerhans of the pancreas. The ability to generate insulin is impaired and patients lose the ability to properly regulate blood glucose levels. Approximately 1.25 million Americans have been diagnosed with T1D, with an expected increase to 5 million by

[1]Critical Path Institute, Tucson, Arizona, USA. *Correspondence: Jagdeep T. Podichetty (jpodichetty@c-path.org)

Received November 9, 2021; accepted February 1, 2022. doi:10.1002/cpt.2559

2050.[1,2] Europe is seeing a steep rise in the number of children and young individuals with T1D in countries with previously lower incidence rates, such as Hungary and Poland.[3–5] Worldwide, T1D has an incidence rate of ~ 15 per 100,000 individuals[6] and T1D-related healthcare cost and lost income amounts to over 16 billion dollars annually.[7]

Insulin replacement therapy remains the mainstay of treatment for T1D and is used to manage blood glucose levels, as the major complications associated with diabetes, such as diabetic ketoacidosis, severe hypoglycemia, nephropathy, neuropathy, retinopathy, and cardiomyopathy, arise from poor short- and long-term glycemic control.

Lowering hemoglobin A1c (HbA1c) levels through intensive insulin therapy reduces the risk of diabetes-related complications,[8] but most patients fail to achieve glycemic targets.[9] Insulin and other blood glucose management approaches have provided substantial benefits to patients,[8] but insulin replacement is not physiologic leaving patients and their caregivers with substantial disease management burden and unmet needs. It also fails to address the underlying etiology of T1D (i.e., autoimmune destruction of pancreatic β-cells).

Given the progressive nature of T1D a staging paradigm was established in 2015 identifying three distinct stages of the T1D disease continuum. Stage 1 marks activation of the autoimmune process, and is characterized by the presence of two or more islet autoantibodies (AAs) with normoglycemia and the absence of clinical symptoms. Stage 2 is defined as the presence of β-cell autoimmunity with dysglycemia and the absence of clinical symptoms, and indicates further β-cell destruction. Finally, stage 3 is defined as the onset of symptomatic disease. This staging paradigm is now widely accepted and is able to support the development of novel therapies aiming to delay or prevent the onset of T1D. Targeting earlier stages of disease is important, as intercepting progression after T1D diagnosis (i.e., preserving residual β-cell function) might be quite late to impart a large clinical benefit, as substantial mass and functionality of endogenous β-cells has already been lost.

Development of therapies that prevent or delay T1D remains challenging. One major barrier lies in quantifying the time-varying risk of progression to stage 3 disease, which can support the design of clinical trials of more feasible duration. Numerous T1D natural history studies have identified individual risk factors for developing T1D. Individuals who have a first-degree relative (FDR) with T1D or express a specific human leukocyte antigen (HLA) haplotype (HLA-DR3/3, DR4/4, DR3/4, DR3/X [X ≠ 3], DR4/X [X ≠ 4]) are likely to be at a higher risk for developing T1D. Additionally, the presence of multiple islet AAs, specifically insulin/proinsulin autoantibody (IAA), glutamic acid decarboxylase 65 autoantibody (GAD65), insulinoma antigen-2 autoantibody (IA-2), or zinc transporter 8 autoantibody (ZnT8), has been shown to be relevant biomarkers for predicting a clinical diagnosis of T1D.[10] Past findings indicate individuals at risk of developing T1D (FDR and specific HLA haplotype) with two or more islet AAs[11] will eventually lead to the onset of T1D over time, and the rate of conversion to stage 3 T1D[12] becomes higher with an increased number of islet AAs.[13–15]

Integrating findings from past natural history studies into patient enrollment criteria in T1D prevention trials is also challenging, given the variability of the latency phase (i.e., progression from stage 1 or 2 to stage 3 disease).[12] To use islet AAs as enrichment biomarkers in T1D prevention trials, quantitative estimates of timing to T1D diagnosis, based on baseline information that is routinely collected in prevention trials is required. By including islet AA status and other clinically relevant covariates, sponsors can quantify the heterogeneity in the latency phase and optimize the design of clinical trials of appropriate and reasonable size, duration, and cost.

In response to this drug development need, the T1D Consortium (T1DC) was founded by the Critical Path Institute (C-Path) in 2017 as a public-private partnership, including pharmaceutical industry, patient-advocacy organizations, philanthropic organizations, clinical researchers, the National Institutes of Health, and the Food and Drug Administration. The primary aim of this consortium was to support drug development in T1D prevention by obtaining the regulatory qualification of biomarkers for T1D prevention trials. T1DC acquired and curated existing patient-level data from 3 observational studies and evaluated the utility of IAA, GAD65, IA-2, and ZnT8 as biomarkers to enrich subjects for inclusion in T1D prevention trials, using a model-based approach. T1DC explored a wide range of patient characteristics, including demographics, HLA-haplotype, FDR T1D status, blood glucose assessments, C-peptide levels, age-adjusted body mass index (BMI), and the various combinations of islet AA presentation. T1DC developed a time-to-event model to quantify the time-varying probability of reaching a diagnosis of T1D, which could be used to optimize subject enrichment strategies for T1D prevention trials aiming to delay or prevent T1D. Based on this model, in March 2020, the European Medicines Agency (EMA) issued a public Letter of Support for "Islet autoantibodies as enrichment biomarkers for type 1 diabetes prevention studies, through a quantitative disease progression model." The EMA letter of support reiterated the need for collection and sharing of relevant data to "permit timely and robust development and validation of such a model...," stating this would "facilitate development and validation of the proposed quantitative [tool]."

Accordingly, the T1DC leveraged shared data from multiple sources to develop quantitative models that provided evidence for formal regulatory submissions through the qualification of novel methodologies pathway at the EMA.[16] The EMA has now adopted its final qualification opinion, formally endorsing the use of the islet AAs, with additional relevant clinical features, as enrichment biomarkers for use in T1D prevention studies.

METHODS

Data

Patient-level data were obtained from three studies, TrialNet (TN01), The Environmental Determinants of Diabetes in the Young (TEDDY), and Diabetes Auto Immunity Study in the Young (DAISY) studies, to perform the analysis and support the qualification effort.[10,17,18] TN01 and TEDDY datasets were used for model development and DAISY was set aside for external validation. All three studies are observational, with different inclusion criteria and scheduled frequency of follow-up.

Data curation and visualization

To build the analysis set, a subset of common variables from all possible variables in each dataset was constructed. Relevant subject features based

on prior knowledge were chosen for the analysis set. These features included (i) presence of islet AAs (IAA, GAD65, IA-2, and ZnT8) measured as a binary variable of either seropositivity or seronegativity, (ii) blood glucose measurements (0 and 120-minute timepoints of oral glucose tolerance tests (OGTTs)), (iii) HbA1c measurements, (iv) demographic information (sex, baseline age, and FDR status), and (v) HLA subtype. The baseline information from the three studies was used for the time-to-event modeling analysis. A derived baseline (**Figure 1**) was used for the analysis set defined as the first record, (i.e., timepoint), for each individual in which the following criteria are satisfied: (i) presence of any two or more islet AAs, (ii) complete, (i.e., non-missing) information for OGTT (0 and 120-minute timepoints), HbA1c measurements, age, and sex. The rationale for choosing two or more islet AAs was based on the utility of the biomarker for drug development, as presence of one or zero islet AAs at baseline have significantly longer expected times to T1D diagnosis compared to two or more islet AAs. These criteria also allow for assessment of glycemic status, and are representative of those likely to enter T1D prevention studies. **Figure S1** shows risk of T1D diagnosis stratified by using only the number of islet AAs present at the first patient record, including zero.

Table S1 lists the baseline covariates evaluated and tested as predictors in the time-to-event model. Of note, C-peptide and BMI were not included as covariates due to significant missing information in the available data. BMI was tested as a covariate in the model and did not show statistical significance as a model predictor. The islet AA combinations were represented separately as binary predictors (presence or absence). As the definition of the derived baseline does not include individuals with missing information for OGTT (0 and 120-minute timepoints), HbA1c, sex, and age. Hence, no imputations were necessary for these variables. All continuous covariates were standardized, (i.e., computed as (original value – mean(value))/(standard deviation of original values)), and OGTT values were first log transformed. The subscript "s" denotes this standardization. A comprehensive tabulation and visualization of the data contained in the analysis set were performed.

Modeling analysis

The flow chart shown in **Figure 2** provides a workflow for model development, where subsequent steps were executed based on best practices for model building and engagement with EMA's Scientific Advice Working Party (SAWP) as part of the qualification pathway. In brief, the model building process included: (i) building a Cox proportional hazard (PH) model using the analysis dataset as a parsimonious initial step; (ii) testing the PH assumption, and if violated continuing the model building process; (iii) building a parametric accelerated failure time (AFT) model using the analysis dataset; (iv) evaluate model performance with k-fold

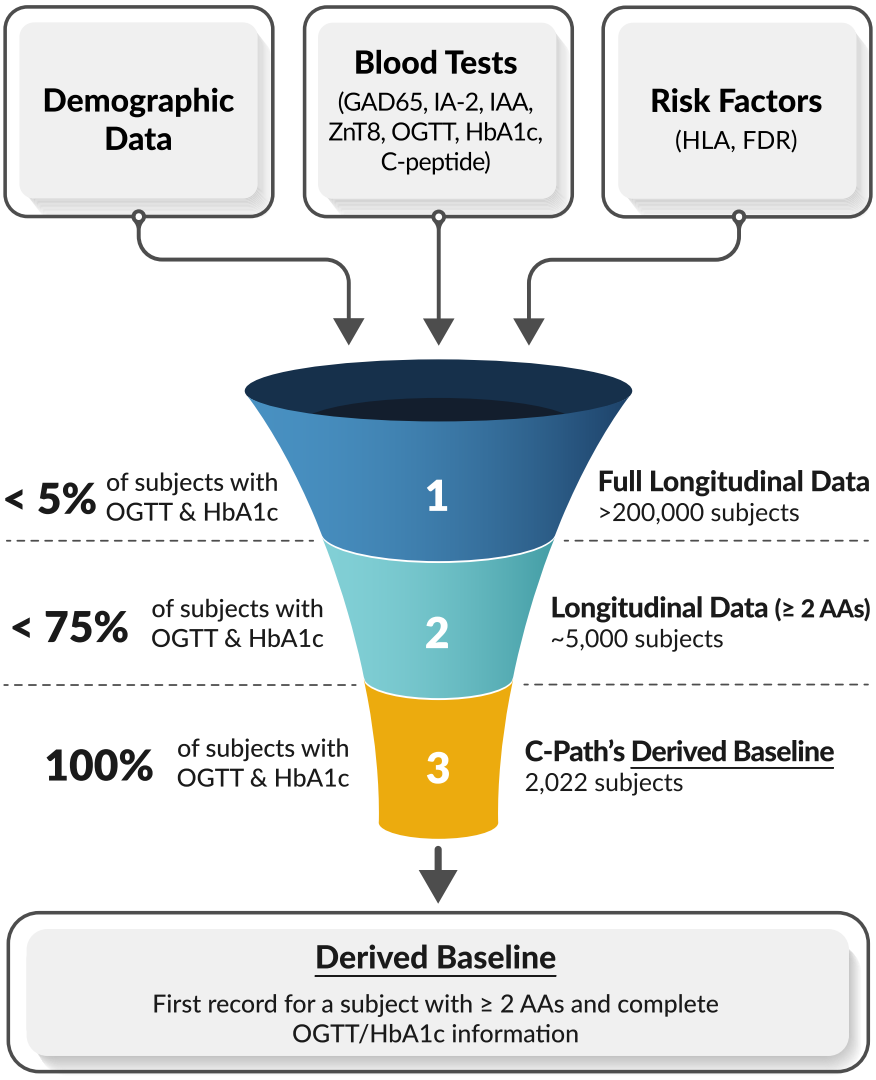

Figure 1 Schematic of data curation process to obtain the derived baseline. OGTT, oral glucose tolerance test. [Colour figure can be viewed at wileyonlinelibrary.com]

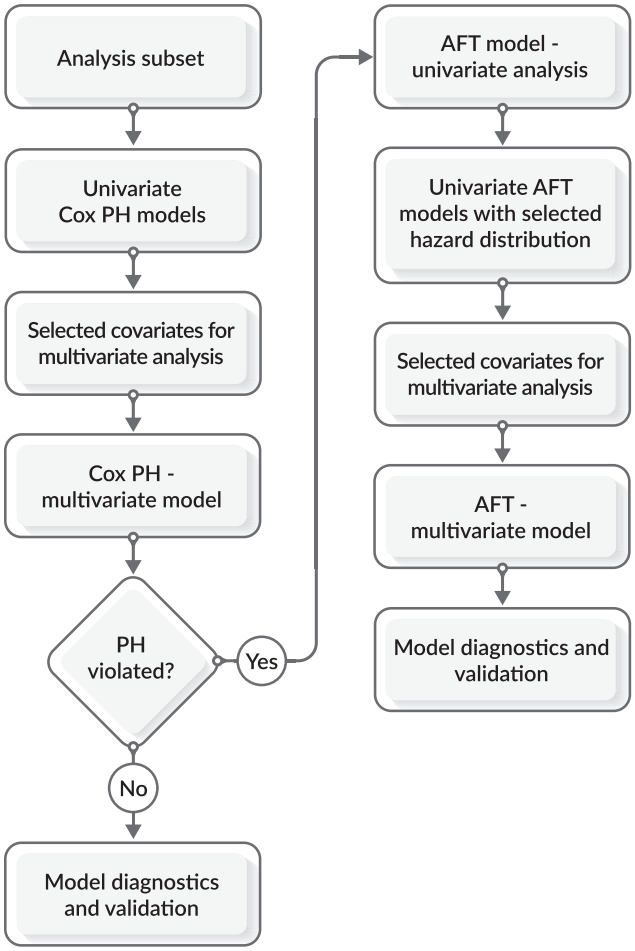

Figure 2 Modeling development workflow. AFT, accelerated failure time; PH, proportional hazard.

cross-validation and external validation using DAISY as an independent dataset. The selection of an AFT methodology was based on initial testing of the PH assumption and consensus with the SAWP regarding the added value of a parametric approach.

The final developed model was an AFT model using a Weibull distribution for the baseline hazard, and predictors of T1D diagnosis. The hazard function for the AFT model is given by:

$$h_i(t) = h_0\left(t/\exp\left(\sum_{j\in I}\beta_j X_{ij}\right)\right)\exp\left(\sum_{j\in I}\beta X_{ij}\right)$$

where $h_i(t)$ is hazard function for individual i determined by a set of j covariates $\{X_{ij}\}$ and corresponding (estimated) coefficients $\{\beta_j\}$, t is the survival time, and $h_o(t)$ is the baseline hazard. The corresponding survival function with a Weibull baseline hazard is given by:

$$S_i(t) = \exp\left\{-\left(\frac{t}{\lambda e^{\sum_j \beta_j X_{ij}}}\right)^{\alpha}\right\}$$

where $S_i(t)$ is the survival function for the i^{th} individual at time t, λ is the scale parameter, and a is the shape parameter.

In both model development phases (Cox and AFT), univariate analyses were carried out and covariates with no significant univariate association (P value ≥ 0.1) to a T1D diagnosis were not considered for the full model development. The Wald test was used to compute the P values and check whether the covariate coefficient value was statistically different from zero. Additionally, a multiplicity adjusted alpha value based on Bonferroni correction was used for univariate analysis. The covariates selected after the univariate analysis were analyzed for multicollinearity and associations before performing multivariate analysis. Pearson's correlation was used to test the correlation between continuous covariates and a correlation value above 0.3 was initially chosen as significant. However, in dialogue with the EMA, it was recommended that certain covariates (age and sex) be considered even with a correlation value above 0.3 as it is expected that future populations may have such correlations preserved. In such cases, accuracy of the model could be improved.

The association between continuous and categorical covariates was tested using the Wilcoxon test and the association between categorical covariates was tested using the chi-square test.

In both cases, a P value < 0.001 (multiplicity adjusted) was used as the threshold for significance.

The multivariate analyses were performed on the covariates selected from the univariate analysis. All possible combinations of covariates were tested in the multivariate analysis as the number of covariates was reasonable. Models were compared using Akaike's information criteria (AIC), and a reduction in AIC value greater than or equal to 10 was considered strong evidence in favor of the model with lower AIC.[19] In the case of the Cox model, the PH assumption was tested using Schoenfeld residuals, which quantify potential time-dependency on survival times. No additional model diagnostics were performed for the Cox PH model due to a violation of the PH assumption observed with the Schoenfeld residuals test.

To build the parametric AFT model, multiple parametric distributions were tested to define the underlying hazard function, including exponential, Weibull, gamma, generalized gamma, generalized F, log logistic, log normal, and Gompertz. AIC values and graphical methods were used to compare the different parametric forms. The "flexsurvreg" function in the "flexsurv" R package was used to perform the selection of the best parametric distribution for the underlying hazard function. The univariate analysis, correlation, and association between covariates analysis, and multivariate analysis for AFT model were performed similar to Cox PH model, as previously described. Model diagnostics were performed using Q-Q plots to test the AFT model assumption for two groups of survival data. In this case, only categorical groupings are permitted, and continuous covariates were split into binary groups according to a chosen threshold.

Model performance and validation

Time-dependent receiver operating characteristic (ROC) curves were generated to assess the predictive performance of the analysis set.[20] For model validation, a k-fold cross-validation technique was used,[21] and data were split into k = 5 subsets with roughly equal numbers of subjects. Four of the five subsets were used as a training set, and the remaining set was used as an individual test set. This process was repeated by assigning one of the five subsets as the new test set, whereas the remaining were used as the training set for all combinations. Additionally, goodness-of-fit plots were generated for all 5 folds and the concordance index (c-index) was computed for each of the 5 folds estimated by time increments of 1 year up to 6 years.

An additional internal validation was performed by analyzing predictive performance on pediatric subpopulations in the data. T1DC selected a random portion (50%) of individuals aged less than 12 years old and used this population as a test data set. The remaining data were used for model training. Goodness-of-fit plots were created by overlaying model estimated survival on Kaplan-Meier curves. The concordance index was computed for time increments of 1 year up to 6 years.

The DAISY dataset was used to perform the external validation. The definition of the derived baseline was applied to obtain the external validation set. The AFT model built using the analysis set was used to predict survival for the external validation set. The visual predictive check (VPC) style goodness-of-fit plots were created to assess the performance of the AFT model on the external dataset.

All analysis was carried out in the R programming language. Model building, visualization, model assumptions, diagnostics, and external

validation was conducted in R (version 4.0.0; R Core Team, Vienna, Austria, 2018) using the packages "survival,"[22] "flexsurv,"[23] "survminer,"[24] "dplyr,"[25] "survAUC,"[26] "rms,"[27] survParamSim,[28] and "riskRegression."[29]

RESULTS

Data curation and visualization

A total of 2,022 subjects were curated for the analysis set with complete information for islet AA positivity, age, sex, HbA1c, and 0 and 120-minute time points of OGTTs. Data summaries for the covariates and the diagnosis information for both TN-01 and TEDDY are shown in **Table S2**. The distribution of diagnosis by each of the continuous covariates is shown in **Figure S2**.

Modeling analysis

As indicated in **Figure 2**, a Cox PH model was first explored. Details can be found in the Supplementary Materials below **Figure S2**. The development of the AFT model was initiated by selecting the most appropriate distribution of the hazard function. Several parametric distributions were tested and compared based on AIC and graphical inspection. The Weibull distribution was found to be the most appropriate parametric distribution.

The univariate analysis using AFT model with Weibull distribution showed covariates age at derived baseline, sex, several islet AA combinations (GAD65_IAA, GAD65_ZnT8, IA-2_ZnT8, IA-2_IAA_ZnT8, and GAD65_IA-2_IAA_ZnT8), log transformed 0 and 120-minute glucose from OGTTs, and HbA1c had statistically significant beta coefficients. The covariates trial ID, BMI, high-risk HLA subtype, FDR, and several AA combinations (GAD65_IA-2, IA-2_IAA, IAA_ZnT8, GAD65_IAA_ZnT8, GAD65_IA-2_ZnT8, and GAD65_IAA_IA-2) did not show a significant effect on overall survival and were dropped from subsequent analysis. Based on the univariate analysis and analysis of correlation and association previously performed for the Cox PH model (see **Supplementary Material**), the covariates GAD65_IAA, GAD65_ZnT8, IA-2_ZnT8, IA-2_IAA_ZnT8, GAD65_IA-2_IAA_ZnT8, Log_GLU0_s, Log_GLU120_s, and HbA1c_s were chosen for AFT multivariate analysis. In the multivariate analysis, a total of eight possible models were obtained with islet AA combinations as the base model (**Table 1**). The models were compared using their AIC values. Model 6 was selected as it produced a lower AIC with fewer covariates. **Table S3** shows shape and scale parameters for the Weibull distribution, estimated beta values, and Wald test *P* values for each covariate.

As stated earlier, based on feedback received from the SAWP, alternative models were developed by including bAGE_s and SEX in different combinations using model 6 (**Table 2**). Although these covariates are strongly associated with others, they were included given the expectation that future populations will have such associations preserved.

For comparison, T1DC will refer to model 6 as the original model (orig_model). The AIC value of alternative model 3 (alt_mod3) was significantly lower (with a reduction > 10) compared with all other alternative models and the original model. Hence, alternative model 3 (alt_mod3) was chosen as the final multivariate parametric AFT model. **Table 3** shows shape and scale parameters for the Weibull distribution, estimated beta values, Wald test *P* values for each covariate, and interpretation of the beta coefficient.

Table 1 Values of AIC for AFT models fitted with a Weibull distribution

Model	Covariates	AIC
1	GAD65_IAA + GAD65_ZnT8 + IA-2_ ZnT8 + IA-2_IAA_ZnT8 + GAD65_IA-2_ IAA_ZnT8 (Base model)	3,292.476
2	Base model + Log_GLU0_s	3,278.769
3	Base model + HbA1c_s	3,173.157
4	Base model + Log_GLU120_s	3,059.067
5	Base model + Log_GLU120_s + Log_GLU0_s	3,052.591
6	Base model + Log_GLU120_s + HbA1c_s	2,981.886
7	Base model + Log_GLU0_s + HbA1c_s	3,172.244
8	Base model + Log_GLU0_s + Log_ GLU120_s + HbA1c_s	2,983.369

AFT, accelerated failure time; AIC, Akaike's information criteria.

The model diagnostics using Q-Q plots indicated that the AFT models adequately described the effect of AA combination status (presence or absence), HbA1c above or below a defined threshold of 5.25%, age as a binary measure above or below 11 years, sex, and the log transformed and standardized 120-minute results from OGTTs as a binary with a threshold of 100 mg/dL (**Figure S3**).

Model performance and validation

The time-dependent ROC analysis and area under the curve (AUC) values showed good predictive performance (**Figure S4**). Additionally, the results showed AUC values greater than 0.8 for up to 2.5 years, which is typically used as a reasonable duration for a clinical trial.

The k-fold cross validation with k = 5 showed c-index values close to or greater than 0.8, indicating good predictive performance. VPC-style plots showed good graphical fit for folds 1, 2, 3, and 4, whereas fold 5 only performed well within the first year. The black curve represents the Kaplan–Meier estimate, and the red curve represents model prediction (**Figure 3**). The internal cross-validation on the pediatric population (age < 12 years old) was derived in the analysis dataset comprised of 1,330 subjects, with 345 from TEDDY and 985 from TN01. Half of this population, (n = 665), were randomly selected as a test set for this cross-validation analysis. The c-index was 0.8 or greater until 3 years and more than 0.75 until 6 years. The VPC style on the pediatric population showed reasonable graphical fit (**Figure S5**).

The external validation performed using the DAISY dataset gave a c-index value of 0.91 in year 1 and 0.82 in year 2, even with the limited number of subjects. The c-index post-3 years were lower compared to the first 2 years, likely due to the sparsity of

Table 2 Value of AIC for original model (model 6) and other alternative models

Model	Covariates	AIC
Original model (orig_mod)	GAD65_IAA + GAD65_ZnT8 + IA-2_ZnT8 + IA-2_IAA_ZnT8 + GAD65_IA-2_IAA_ZnT8+ Log_GLU120_s + HbA1c_s	2,982
Alternative model 1 (alt_mod1)	GAD65_IAA + GAD65_ZnT8 + IA-2_ZnT8 + IA-2_IAA_ZnT8 + GAD65_IA-2_IAA_ZnT8+ Log_GLU120_s + HbA1c_s + SEX	2,972
Alternative model 2 (alt_mod2)	GAD65_IAA + GAD65_ZnT8 + IA-2_ZnT8 + IA-2_IAA_ZnT8 + GAD65_IA-2_IAA_ZnT8+ Log_GLU120_s + HbA1c_s + bAGE_s	2,937
Alternative model 3 (alt_mod3)	GAD65_IAA + GAD65_ZnT8 + IA-2_ZnT8 + IA-2_IAA_ZnT8 + GAD65_IA-2_IAA_ZnT8+ Log_GLU120_s + HbA1c_s + bAGE_s + SEX	2,921

AIC, Akaike's information criteria.

Table 3 Final selected model (alt_mod3) parameter estimates

Covariates	Beta	95% lower CI	95% upper CI	*P* value	Interpretation of beta coefficient
Shape	1.370	1.280	1.470	< 0.0001	Weibull Shape parameter
Scale	6.780	5.990	7.670	< 0.0001	Weibull Scale parameter
log_GLU120_s	−0.546	−0.623	−0.469	< 0.0001	Unit increase in log_GLU120_s value reduces the time to T1D diagnosis by 40%
HbA1c_s	−0.322	−0.392	−0.252	< 0.0001	Unit increase in HbA1c_s value reduces the time to T1D diagnosis 30%
SEX	0.275	0.147	0.403	< 0.0001	Having SEX = Male increases the time to T1D Diagnosis by 30%
bAGE_s	0.267	0.183	0.350	< 0.0001	Unit increase in bAGE_s value increases the time to T1D diagnosis by 30%
GAD65_IAA	0.506	0.284	0.728	< 0.0001	Presence of GAD65_IAA increases the time to T1D diagnosis 70%
GAD65_ZnT8	0.474	0.225	0.723	0.0002	Presence of GAD65_ZnT8 increases the time to T1D diagnosis by 60%
IA-2_ZnT8	−0.346	−0.603	−0.087	0.0084	Presence of IA-2_ZnT8 reduces the time to T1D diagnosis by 30%
IA-2_IAA_ZnT8	−0.257	−0.512	−0.002	0.0482	Presence of IA-2_IAA_ZnT8 reduces the time to T1D diagnosis by 20%
GAD65_IA-2_IAA_ZnT8	−0.064	−0.226	0.099	0.4400	Presence of GAD65_IA-2_IAA_ZnT8 reduces the time to T1D diagnosis by 10%

CI, confidence interval; IA, insulin antibody; IAA, insulin/proinsulin autoantibody; T1D, type 2 diabetes.

T1D diagnoses during the later years in the DAISY dataset. The VPC plot showed good graphical fit even with limited number of events (**Figure 4**). These findings provide evidence for good predictive performance of the model for time duration over which a trial of reasonable duration would be conducted.

DISCUSSION

This work aimed to leverage existing data sources that captured islet AA measurements and glycemic markers in a population likely to participate in T1D prevention trials and generate evidence supporting an EMA qualification opinion for the islet AAs as enrichment biomarkers in T1D prevention trials. Data were obtained from observational studies and harmonized to ensure interoperability. Initially, a semiparametric Cox PH model was assessed using an analysis set from the curated dataset. During model analysis, the PH assumption was found to not hold true. Given SAWP's recommendation to explore a parametric modeling approach, a parametric AFT model was selected, as this does not require the PH assumption to hold true.

In the model, baseline islet AA positivity was represented as a single covariate with 11 distinct levels, representing all possible combinations of 2 or more relevant islet AAs. This approach is more comprehensive than considering the total numbers of islet AAs and allows for quantification of risk by islet AA type. Performing a cross-sectional assessment permits the possibility that subjects positive for two or three islet AAs may convert to the three or four islet AAs before diagnosis. However, this method reflects how sponsors will recruit subjects for T1D prevention studies, as subject's islet AA time history will likely not be available to sponsors. Hence, the use of baseline information is preferred in this context.

Results from the AFT modeling analysis showed GAD65_IAA and GAD65_ZnT8 combinations have the least relative risk compared with all other combinations, whereas IA-2_ZnT8 has the highest relative risk. Presence of any three islet AAs was

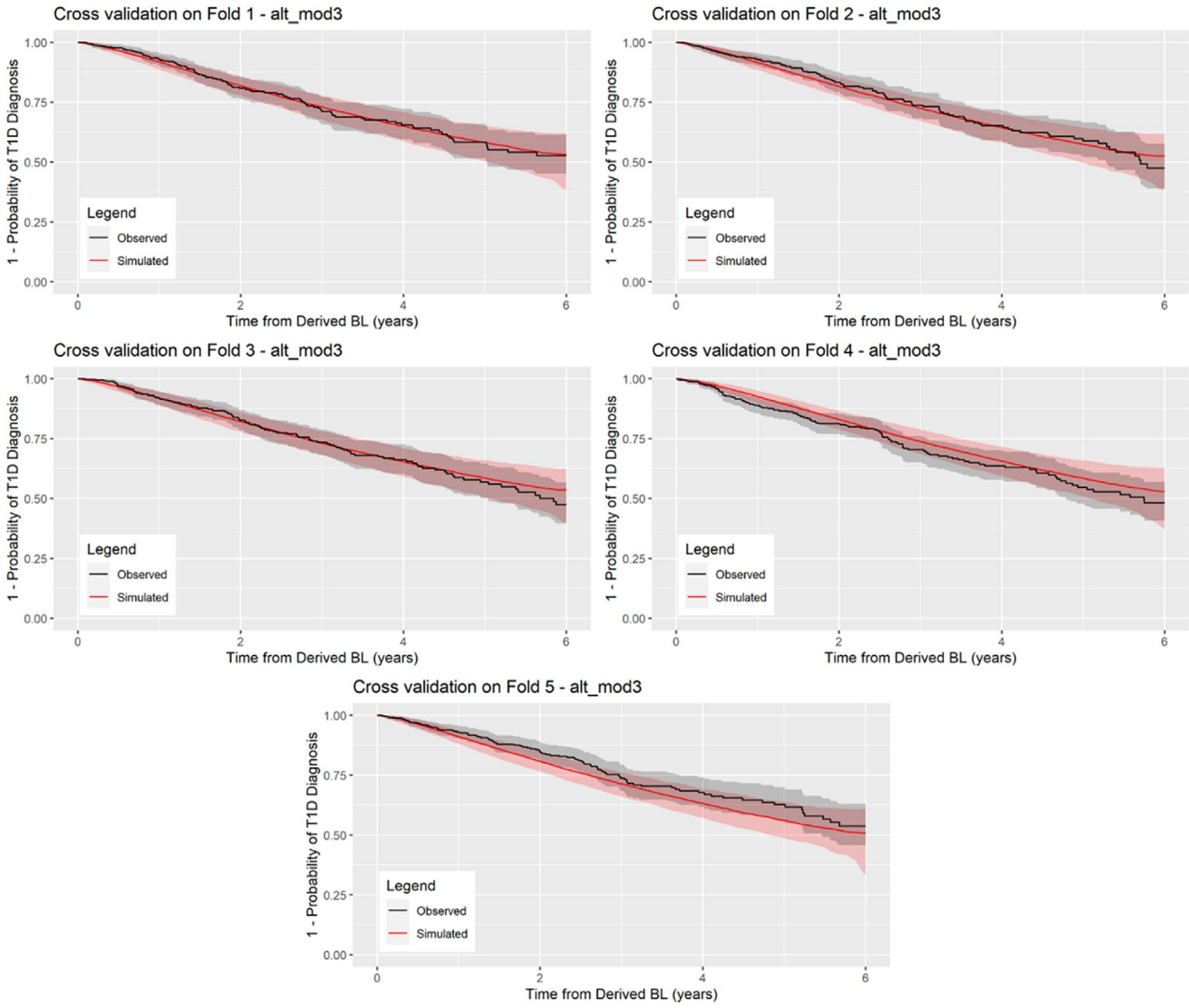

Figure 3 Visual predictive check (VPC)-style plots for k-fold cross validation (red shaded region shows the 95% prediction interval and the black shaded region shows the 95% confidence interval (CI) for the observed data). [Colour figure can be viewed at wileyonlinelibrary.com]

not shown to be significantly different from baseline hazard, and presence of all four islet AAs had a marginal risk increase relative to the baseline hazard. Incorporating 120-minute OGTTs, baseline age, sex, and HbA1c values provide a significant ability to further stratify risk of T1D diagnosis within this islet AA positive populations.

Both internal and external validation procedures were carried out for the final selected model (alt_mod3). During internal validation, a time-dependent ROC analysis showed high overall concordance across AUC values (> 0.75), especially within the first 2-years following the derived baseline, which represents a time frame concordant with feasible trial design for T1D prevention. Concordance in the first 2 years was also high (c-index > 0.75). As pediatric populations (< 12 years of age) are of keen interest to sponsors, additional internal cross-validation was carried out in this population, also resulting in a high degree of concordance (c-index ≈ 0.8). External validation using the DAISY study showed a high concordance in the first 2 years (c-index > 0.8) in this population.

An important data consideration was the selection of individuals with non-missing glycemic measurement information for inclusion in the derived baseline population used in the AFT model. This population is representative of those likely to enter a T1D prevention study. The modeling analysis indicated that dysglycemia measured by OGTT is highly predictive of timing to T1D diagnosis, thereby providing utility for further stratification. Another key data consideration is the unknown time history of islet AA positivity in TN01, representing a source of variability. Although this variability is contained by each islet AA combination, it represents the practical reality of drug development and trial design for prevention studies. The purpose of this effort was to qualify islet AAs as enrichment biomarkers for T1D prevention studies. As such, study sponsors will typically not know the islet AA time history of participating subjects, and the TN01 data used are therefore representative of a population likely to enter T1D prevention studies. Finally, the size of the external validation was relatively small, due to the definition of

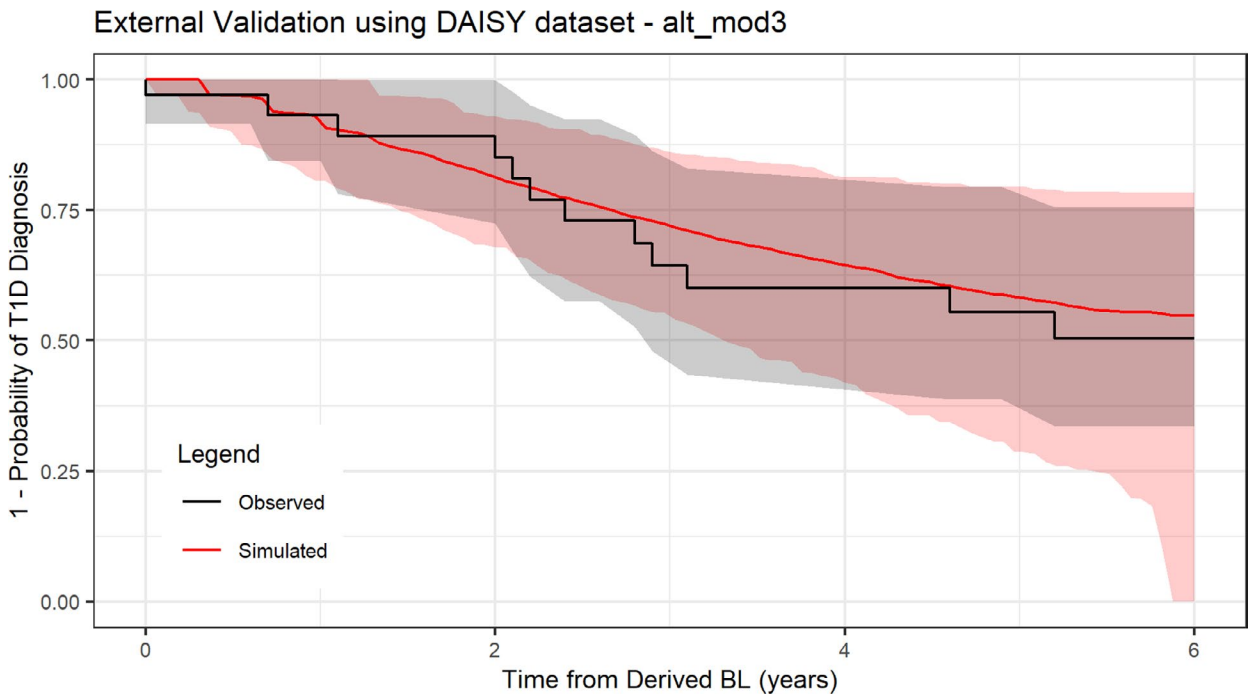

Figure 4 Visual predictive check (VPC)-style plot for external validation using the Diabetes Auto Immunity Study in the Young (DAISY) analysis dataset (red shaded region shows the 95% prediction interval and the black shaded region shows the 95% confidence interval (CI) for the observed data). BL, baseline. [Colour figure can be viewed at wileyonlinelibrary.com]

baseline used in the analysis. Although the model performed well, additional credibility could be established with larger numbers of subjects from other independent datasets.

In conclusion, T1DC's analysis of integrated data from independent observational data sources represented adequate supporting evidence to receive EMA qualification for the use of islet AAs as enrichment biomarkers for T1D prevention trials. When used in this setting, islet AAs can identify populations likely to reach a T1D diagnosis during T1D prevention studies of reasonable duration. The model presented provides a basis to quantitatively link independent sources of risk, measured by islet AAs, baseline age, sex, and glycemic measures. Regulatory endorsement of the islet AAs (when used according to the qualified context-of-use) through the EMA qualification of novel methodologies pathway, based on the evidence provided through the modeling exercise, is expected to facilitate drug development in T1D prevention by providing sponsors with increased confidence and regulatory certainty when implementing these biomarkers into clinical trial design. The underpinning model will support a fully functioning end-user tool enabling sponsors to optimize enrichment criteria for T1D prevention studies. In accordance with T1DC's and EMA's qualification of novel methodologies goals, the model will be made publicly available.

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).

ACKNOWLEDGMENTS

The authors would like to thank the T1DC members for their guidance and support. The TEDDY and TN01 studies were conducted by the TEDDY and TN01 Investigators and supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). The data from TEDDY and TN01 reported here were supplied by the NIDDK Central Repositories. This manuscript was not prepared in collaboration with Investigators of the TEDDY and TN01 studies and does not necessarily reflect the opinions or views of the TEDDY and TN01 studies, the NIDDK Central Repositories, or the NIDDK. The T1DC recognizes the DAISY study group for their data contribution and support. The T1DC thanks Alistair J. K. Williams and Liping Yu for their contributions to the documentation of the islet AA assays used in the TEDDY, TN01, and DAISY studies.

FUNDING

The T1DC and this project are supported by funds from JDRF International, Helmsley Charitable Trust, Janssen Research and Development, LLC, Novo Nordisk, and ProventionBio. Critical Path Institute is supported by the Food and Drug Administration (FDA) of the U.S. Department of Health and Human Services (HHS) and is 56.5% funded by FDA/HHS, totaling $16,749,891, and 43.5% funded by non-government source(s), totaling $12,895,366. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement by, the FDA/HHS or the U.S. Government. For more information, please visit FDA.gov.

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS

J.T.P., I.M.D., S.E.D., S.R.K., J.K.B., and K.R. wrote the manuscript. J.T.P., I.M.D., R.N.M., K.R., and J.K.B. designed the research. J.T.P., P.L., R.N.M., and J.K.B. performed the research. P.L. and L.S.S. analyzed the data.

1. Type 1 Diabetes Statistics. Beyond Type 1. <https://beyondtype1.org/type-1-diabetes-statistics/>. Accessed October 23, 2019.
2. Dabelea, D. *et al*. Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. *JAMA* **311**, 1778–1786 (2014).
3. Barkai, L. *et al*. Changes in the incidence and prevalence of type 1 and type 2 diabetes among 2 million children and adolescents in Hungary between 2001 and 2016 – a nationwide population-based study. *Arch. Med. Sci.* **16**, 34–41 (2019).
4. Skrivarhaug, T., Stene, L.C., Drivvoll, A.K., Strøm, H., Joner, G., & Norwegian Childhood Diabetes Study Group. Incidence of type 1 diabetes in Norway among children aged 0–14 years between 1989 and 2012: has the incidence stopped rising? Results from the Norwegian Childhood Diabetes Registry. *Diabetologia* **57**, 57–62 (2014).
5. Szalecki, M. *et al*. Epidemiology of type 1 diabetes in Polish children: a multicentre cohort study. *Diabetes Metab. Res. Rev.* **34** (2018).
6. Mobasseri, M., Shirmohammadi, M., Amiri, T., Vahed, N., Hosseini Fard, H. & Ghojazadeh, M. Prevalence and incidence of type 1 diabetes in the world: a systematic review and meta-analysis. *Health Promot. Perspect.* **10**, 98–115 (2020).
7. Association. Economic costs of diabetes in the U.S. in 2017. *Diabetes Care* **41**, 917–928 (2018).
8. Writing Group for the DCCT/EDIC Research Group *et al*. Association between 7 years of intensive treatment of type 1 diabetes and long-term mortality. *JAMA* **313**, 45–53 (2015).
9. Foster, N.C. *et al*. State of type 1 diabetes management and outcomes from the T1D exchange in 2016–2018. *Diabetes Technol. Ther.* **21**, 66–72 (2019).
10. Hagopian, W.A. *et al*. The Environmental Determinants of Diabetes in the Young (TEDDY): genetic criteria and international diabetes risk screening of 421 000 infants. *Pediatr. Diabetes* **12**, 733–743 (2011).
11. Ziegler, A.G. *et al*. Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. *JAMA* **309**, 2473 (2013).
12. Insel, R.A. *et al*. Staging presymptomatic type 1 diabetes: a scientific statement of JDRF, the endocrine society, and the American Diabetes Association. *Diabetes Care* **38**, 1964–1974 (2015).
13. Sosenko, J.M. *et al*. Incident dysglycemia and progression to type 1 diabetes among participants in the Diabetes Prevention Trial-Type 1. *Diabetes Care* **32**, 1603–1607 (2009).
14. Veijola, R., Koskinen, M., Helminen, O. & Hekkala, A. Dysregulation of glucose metabolism in preclinical type 1 diabetes. *Pediatr. Diabetes* **17**(Suppl 22), 25–30 (2016).
15. Xu, P., Krischer, J.P., & Type 1 Diabetes TrialNet Study Group. Prognostic classification factors associated with development of multiple autoantibodies, dysglycemia, and type 1 diabetes-A recursive partitioning analysis. *Diabetes Care* **39**, 1036–1044 (2016).
16. European Medicines Agency Opinions and letters of support on the qualification novel methodologies for medicine development. European Medicines Agency. Published 2021. <https://www.ema.europa.eu/en/human-regulatory/research-development/scientific-advice-protocol-assistance/novel-methodologies-biomarkers/opinions-letters-support-qualification-novel-methodologies-medicine-development>. Accessed July 14, 2021.
17. Mahon, J.L. *et al*. The TrialNet natural history study of the development of type 1 diabetes: objectives, design, and initial results. *Pediatr. Diabetes* **10**, 97–104 (2009).
18. Rewers, M. *et al*. Newborn screening for HLA markers associated with IDDM: diabetes autoimmunity study in the young (DAISY). *Diabetologia* **39**, 807–812 (1996).
19. Burnham, K.P. & Anderson, D.R. Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Methods Res.* **33**, 261–304 (2004).
20. Heagerty, P.J. & Zheng, Y. Survival model predictive accuracy and ROC curves. *Biometrics* **61**, 92–105 (2005).
21. Breiman, L. & Spector, P. Submodel selection and evaluation in regression. The X-random case. *Int. Stat. Rev.* **60**, 291 (1992).
22. Therneau TM. Package 'survival' Version 3.1-12. Published online April 10, 2020 <https://cran.r-project.org/web/packages/survival/survival.pdf>.
23. Jackson, C.H. flexsurv: a platform for parametric survival modeling in R. *J. Stat. Softw.* **70**, i08 (2016).
24. Kassambara, A., Kosinski, M. survminer: Drawing Survival Curves using 'ggplot2'. *URL HttpsCRAN R-Proj Orgpackage Survminer R Package Version 04* **3**, 15 (2018).
25. Wickham, H., François, R., Henry, L., Müller, K. & RStudio. Dplyr: a grammar of data manipulation. 2020. <https://CRAN.R-project.org/package=dplyr>. Accessed April 21, 2020.
26. Potapov, S., Adler, W. & Schmid, M. Package 'SurvAUC', estimators of prediction accuracy for time-to-event data (2015). https://www.r-project.org/conferences/useR-2011/abstracts/310311-matthiasschmid.pdf.
27. Harrell, F.E. *Rms:Regression Modeling Strategies*. 2019. <https://CRAN.R-project.org/package=rms>. Accessed April 21, 2020.
28. Yoshida, K. & Claret, L. *SurvParamSim:Parametric Survival Simulation with Parameter Uncertainty*. 2020. <https://CRAN.R-project.org/package=survParamSim>. Accessed August 24, 2020.
29. Ozenne, B., Lyngholm, S.A., Scheike, T., Torp-Pedersen, C. & Alexander, G.T. riskRegression: predicting the risk of an event using cox regression models. *R J.* **9**, 440 (2017).