Written Response to Issues 1 – 3:

This document contains written responses to EMA's Qualification Opinion list of issues 1 through 3, as per the request from the Scientific Advice Working Party (SAWP). The responses presented below represent corresponding modifications in the Briefing Dossier.

Issue 1: The statistical notation and the description of the model are incorrect. The applicant suggests that covariates that are introduced in the model influence the baseline hazard (h0), this is only partially correct. The R-package that the applicant is using to fit the survival models clearly states: "By default, covariates are placed on the "location" parameter of the distribution, typically the "scale" or "rate" parameter, through a linear model (…)". This means that the covariates in the model only influence the "scale" parameter (λ) of the Weibull distribution and not the "shape" parameter (α). The statistical notation in equation 2 on p.29 should be adjusted to better reflect this modelling approach. The applicant is requested to provide the model with the correct description.

Applicant Response: The statistical notation in Equation 2 on p.29, Section 4.3.5.5 of the Briefing Dossier is modified with the following equation to define the Accelerated Time Failure (AFT) model, implemented in the "Flexsurv" R package (p. 30, Section 4.3.4.2) (Jackson 2016) as follows:

$$S_i(t) = S_0\left(t\exp(-\sum_{j\in I}\beta_j X_{ij})\right)$$

Where S_0 is a prespecified form of the parametric distribution for the survival function such as Weibull, Lognormal, Log-logistic, Gamma, and Generalized gamma. **Table 1** provides the survival functions for a list of different forms of parametric distribution.

Table 1. Survival function with various forms of parametric distributions

Parametric Distribution	Survival function	Parameter
Weibull	$S_i(t) = \exp\{-\left(\frac{t}{\lambda e^{\sum_j \beta_j X_{ij}}}\right)^{\alpha}\}$	Where λ is the scale parameter and α is the shape parameter.
Log-normal	$S(t_i) = 1 - \Phi\left(\frac{\ln(T) - \mu - \sum_{j\epsilon I}\beta_j X_{ij}}{\sigma}\right)$	σ is the shape parameter and μ is location
Log-logistic	$S(t_i) = \frac{1}{1 + (\lambda t_i^{\alpha}{}_i e^{-\alpha\sum_j \beta_j X_{ij}})}$	Where λ is the scale parameter and α is the shape parameter.
Gamma	$S(t_i) = \int_t^{\infty} \frac{\lambda^{\alpha}(c(u))^{\alpha-1}e^{-(\lambda c(u))}}{\Gamma(\alpha)}du$ Where $c(u) = u * e^{-\sum_j \beta_j X_{ij}}$	Where λ is the rate parameter and α is the shape parameter.
Generalized gamma	$S(t_i) = \int_t^{\infty} \frac{p\lambda^{p\alpha}(c(u))^{p\alpha-1}e^{-(\lambda c(u))^p}}{\Gamma(\alpha)}du$ Where c(u)= $u * e^{-\sum_j \beta_j X_{ij}}$	Where λ is the rate parameter and p is the shape parameter.

For the R-package survreg, the output "intercept" is the log of the scale (λ) and the output "scale" is the inverse of the shape (α) parameter as shown in **Table 1**. For the R-package flexsurvreg, the output "scale" is the scale (λ) and the output "shape" is the shape (α) parameter as shown in **Table 1**.

Issue 2: Please compare the predictive performance of the proposed model with that of alternative models with other combinations of covariates including a model with baseline age and sex in addition to the covariates identified by the applicant as the final AFT model.

Applicant Response: Alternative models were tested with different combinations of covariates including baseline age and sex in addition to the covariates previously included in the model. The AFT model 6, with the lowest Akaike Information Criterion (AIC) among models initially tested, presented in the previous briefing dossier (Section 4.4.2.4) is referenced here as the original model (orig_mod). **Table 2** shows the selected covariates for the alternative models. The predictive performance for these models was compared using the AIC **(Table 2)**. The AIC value of alternative model 3 (alt_mod3) was significantly lower (with a reduction > 10) compared to all other alternative models and the original model. Hence, alternative model 3 (alt_mod3) was chosen as the selected model. **Table 3** shows the parameter estimates for the selected model (alt_mod3).

Model performance for the selected model (alt_mod3) was assessed using time dependent Receiver Operating Characteristic (ROC) curves and associated area under the curve (AUC) values **(Figure 1)**. The internal validation for the selected model (alt_mod3) was performed using visual predictive check (VPC)-style plots for a k-fold cross-validation and an internal validation with a pediatric population. An external validation was performed with the DAISY dataset **(Figures 2-4)** and c-index values over 6 years **(Tables 4-6)**. The VPC-style plots overlaying observed data over model predictions showed good graphical fit. The "survParamSim" package was used to generate the VPC-style plots for **Figures 2-4**.

The time-dependent ROC curves and AUC values showed good prediction performance especially for up to 2.5 years with AUC values greater than 0.8 (**Figure 1**). The AUC values for subsequent years for up to 5.5 years were greater than 0.75. These results provide evidence for good predictive power for time frames over which clinical trials of reasonable duration would be conducted. The c-index for the selected model (alt_mod3) for all five folds over six years was in most cases close to or higher than 0.8, suggesting good predictive performance **(Table 4)**. VPC-style plots overlaying Kaplan-Meier curves over the selected model predictions showed good graphical fit for folds 1, 2, 3 and 4 while 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 2)**. For the internal cross validation using a pediatric population (age < 12), a c-index of 0.8 or higher was obtained until 3 years and a c-index of 0.75 or higher was obtained up to 6 years for the selected model (alt_mod3) indicating good model performance **(Table 5)**. The visual predictive check (VPC) performed on the survival plot for cross-validation on the pediatric population (age < 12) showed reasonable graphical fit **(Figure 3)**. For external validation with DAISY dataset, the selected model (alt_mod3) achieved a c-index 0.91 and 0.82 in years one and two, respectively, even with a limited number of subjects (n=34) **(Table 6)**. However, the c-index values beyond three years were relatively lower than up to 2 years, likely attributable to the sparsity of T1D diagnoses during the later years in the DAISY analysis set **(Table 6)**. The VPC performed on the survival plot showed good graphical fit **(Figure 4)**. These results

provide strong evidence for good predictive power for time frames over which a trial of reasonable duration would be conducted.

Table 2. Value of AIC for the original model 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	2982
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	2972
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	2937
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	2921

Table 3. Selected model (alt_mod3) parameter estimates

Covariates	Beta	95% lower CI	95% upper CI	p-value
Shape	1.370	1.280	1.470	4.31E-192
Scale	6.780	5.990	7.670	4.36E-56
log_GLU120_s	-0.546	-0.623	-0.469	1.54E-43
HbA1c_s	-0.322	-0.392	-0.252	1.33E-19
SEX	0.275	0.147	0.403	2.65E-05
bAGE_s	0.267	0.183	0.350	3.57E-10
GAD65_IAA	0.506	0.284	0.728	7.95E-06
GAD65_ZnT8	0.474	0.225	0.723	1.88E-04
IA-2_ZnT8	-0.346	-0.603	-0.089	8.42E-03
IA-2_IAA_ZnT8	-0.257	-0.512	-0.002	4.82E-02
GAD65_IA-2_IAA_ZnT8	-0.064	-0.226	0.099	4.40E-01

Figure 1. Evaluation of model performance using time dependent receiver operation characteristic (ROC) analysis on the selected model (alt_mod3). TPR is true positive rate and FPR is false positive rate.

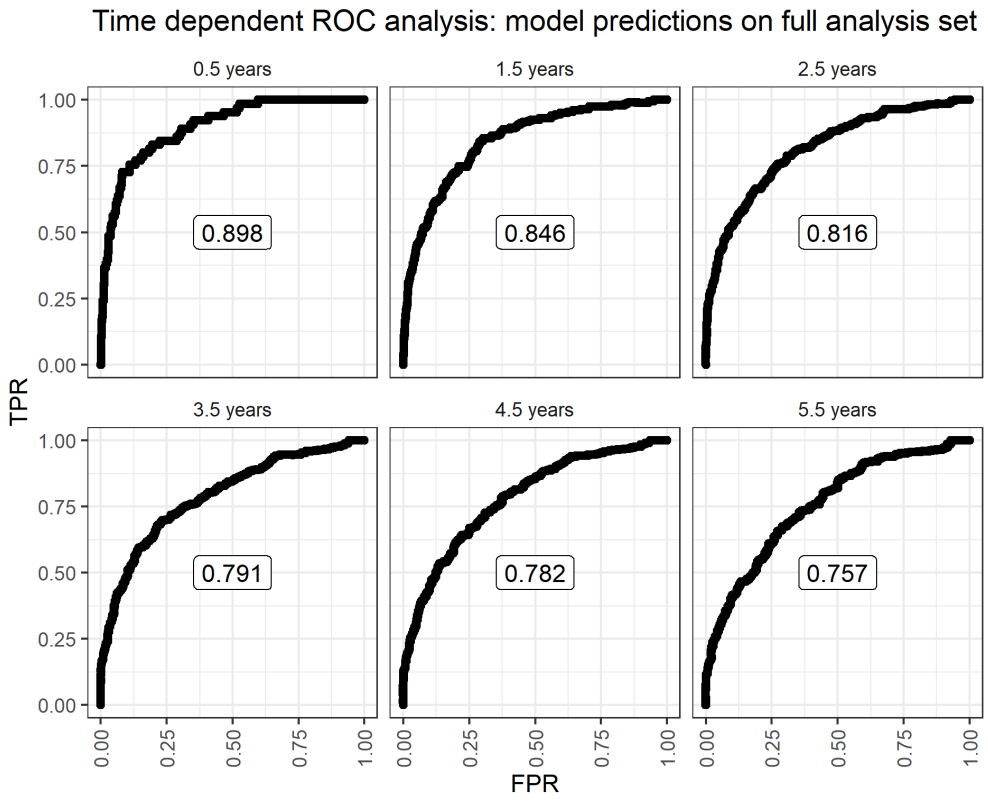


Table 4. Selected model (alt_mod3) model c-index values over 6 years for each fold during k-fold cross validation analysis

C-index (alt_mod3)	Up to year 1	Up to year 2	Up to year 3	Up to year 4	Up to year 5	Up to year 6
fold 1	0.81	0.76	0.75	0.75	0.75	0.74
fold 2	0.87	0.85	0.81	0.81	0.80	0.79
fold 3	0.85	0.82	0.80	0.78	0.77	0.77
fold 4	0.84	0.82	0.80	0.79	0.78	0.78
fold 5	0.87	0.83	0.81	0.81	0.80	0.80

Figure 2. Selected model (alt_mod3) 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 for the observed data). Baseline is referred to as BL in the figures.

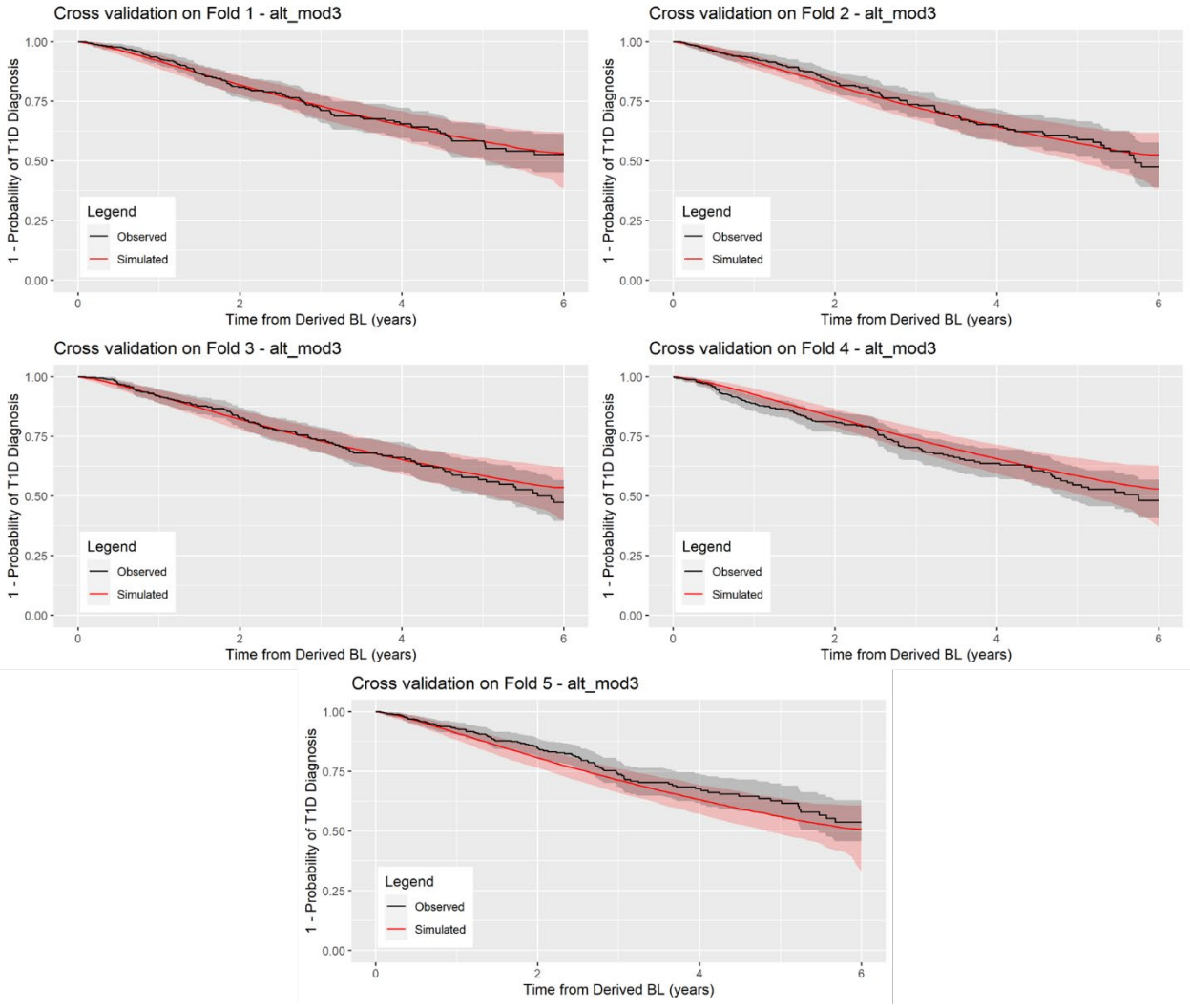


Table 5. C-index values over six years with cross-validation on a pediatric population (age < 12) for the selected model (alt_mod3)

	Up to year 1	Up to year 2	Up to year 3	Up to year 4	Up to year 5	Up to year 6
C-index	0.88	0.84	0.81	0.79	0.78	0.78

Figure 3. Selected model (alt_mod3) VPC-style plot for internal cross validation (CV) using a pediatric population (age < 12 yrs) (red shaded region shows the 95% prediction interval and the black shaded region shows the 95% confidence interval for the observed data). Baseline is referred to as BL in the figure.

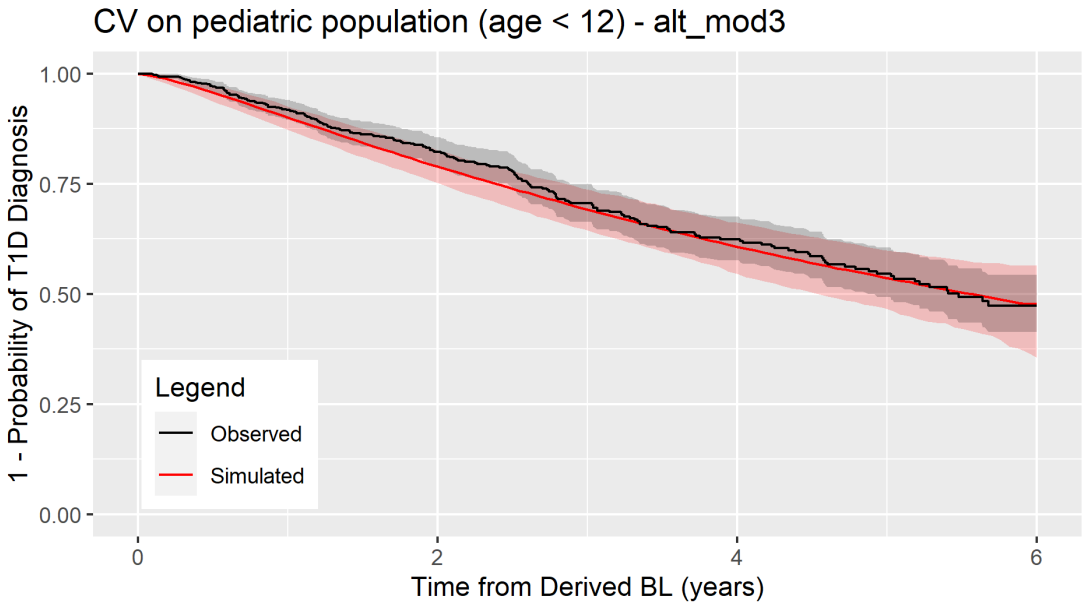
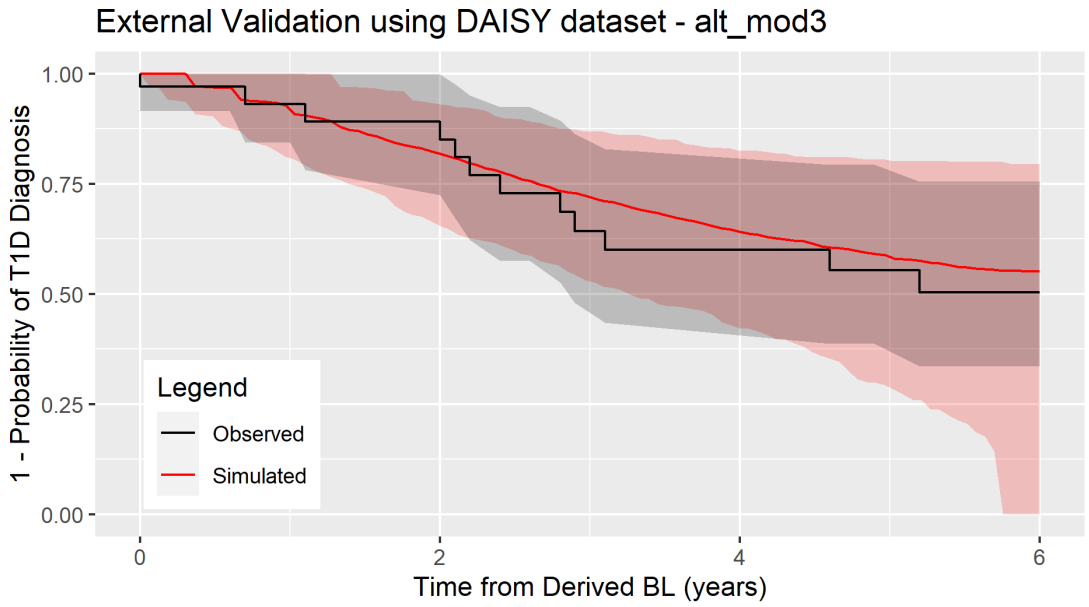


Table 6. C-index values over six years with DAISY external validation dataset for the selected model (alt_mod3)

	Up to year 1	Up to year 2	Up to year 3	Up to year 4	Up to year 5	Up to year 6
C-index	0.91	0.82	0.67	0.68	0.67	0.66

Figure 4. Selected model (alt_mod3) VPC-style plot for external validation using DAISY dataset (red shaded region shows the 95% prediction interval and the black shaded region shows the 95% confidence interval for the observed data).
Baseline is referred to as BL in the figure.



Issue 3: Please provide a "visual predictive check"-style figure instead of Figures 9, 10 and 11, along with the R-code used to generate the VPCs.

Applicant Response: The results in section 4.4 of the Briefing Dossier were updated using the "survParamSim" package to generate the VPC-style plots for Figures 9, 10, and 11. The procedure mimics parametric bootstrap simulations and follows the steps below:

1. Estimate the model using the training set.
2. Using the parameter estimates ($\widehat{\boldsymbol{\beta}}$) and variance-covariance matrix ($\widehat{\boldsymbol{\Sigma}}_{\beta}$)) from the model in step 1, sample parameter values ($\boldsymbol{\beta}$) from multivariate Normal ($\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\Sigma}}_{\beta}$)
3. Generate event T_i^* times using the covariates in the validation set and parameters generated in step 2 from Weibull distribution/Parametric distribution. (For Weibull, the scale= $\lambda e^{\sum_j \beta_j x_{ij}}$ and shape= 1/scale from survreg estimates). Generate censoring times (C_i) as uniform random values.
4. Define simulated event indicator/status as $\delta_i = \mathrm{I}\,(T_i^* \leq C_i)$ and observed event times $T_i = \min\,(T_i^*, C_i)$
5. Derive Kaplan Meier estimates for the simulated sample. Interpolate survival times at smaller ranges i.e. year/event times of validation set. "approx." function in r is used to interpolate
6. Repeat steps 2-5 1000 times. From the 1000 survival estimates in 5, plot 95% predicted intervals at prespecified time points. Overlay Kaplan Meier plot of validation/observed data set.

Figure 9. VPC-style plots for k-fold cross-validation (red shaded region shows the 95% prediction interval and black shaded region shows the 95% confidence interval for the observed data). Baseline is referred to as BL in the figures.

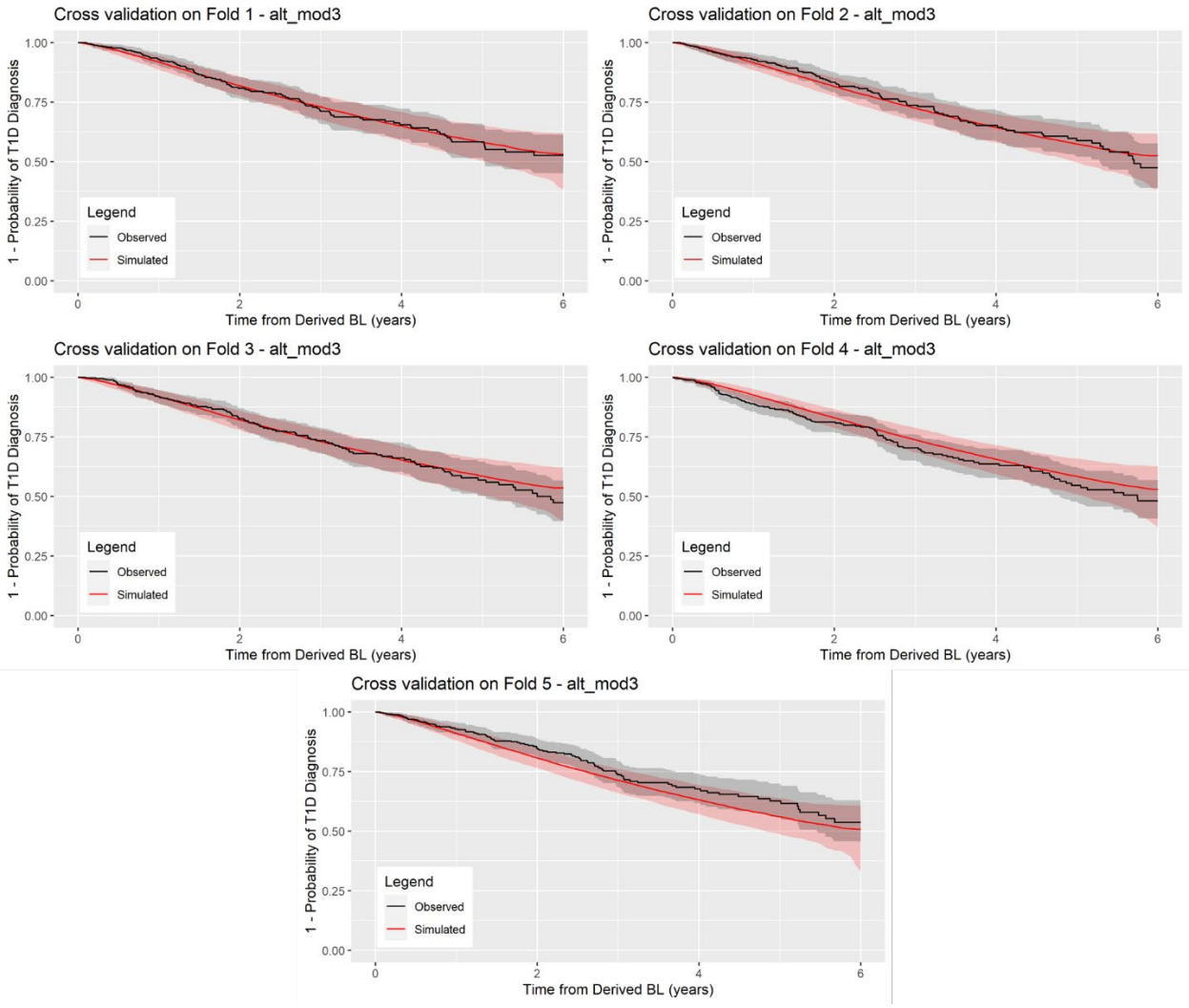


Figure 10. VPC-style plot for internal cross validation (CV) using a pediatric population (age < 12 yrs) (red shaded region shows the 95% prediction interval and the black shaded region shows the 95% confidence interval for the observed data). Baseline is referred to as BL in the figure.

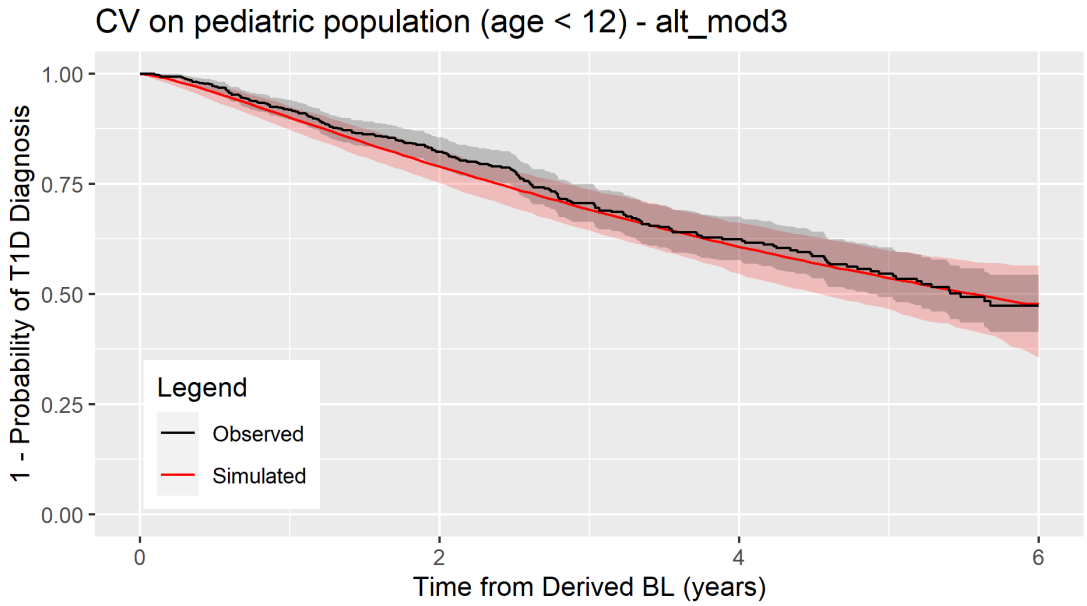
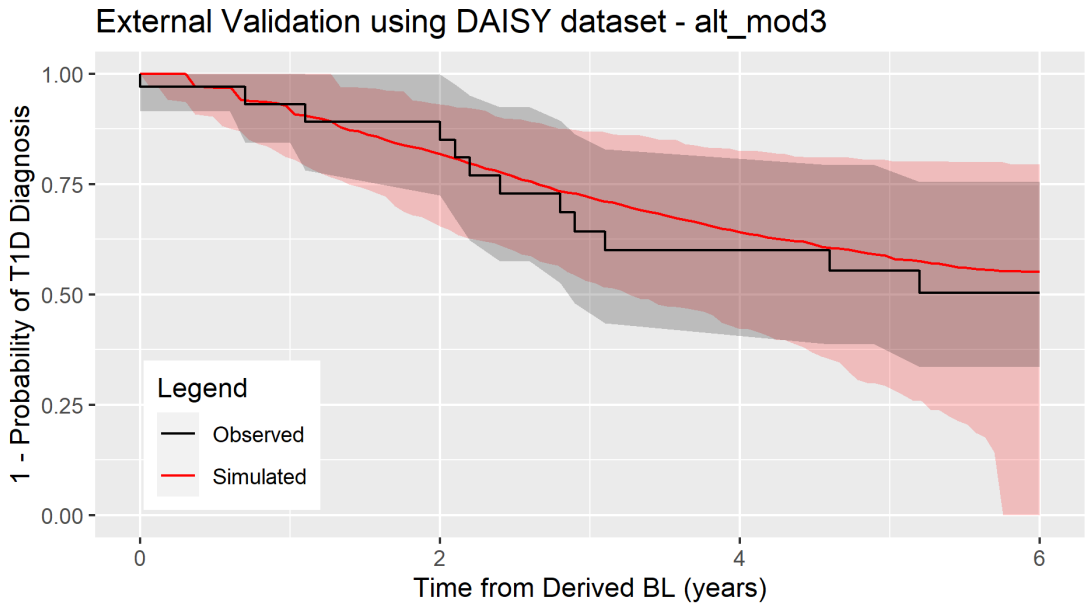


Figure 11. VPC-style plot for external validation using DAISY dataset (red shaded region shows the 95% prediction interval and black shaded region shows the 95% confidence interval for the observed data). Baseline is referred to as BL in the figure.



References

Jackson, Christopher H. 2016. "Flexsurv: A Platform for Parametric Survival Modeling in R." *Journal of Statistical Software* 70 (8). https://doi.org/10.18637/jss.v070.i08.