Critical Path Institute

Type 1 Diabetes Consortium

Modeling & Simulation Team

Islet Autoantibodies as Enrichment Biomarkers for Type 1 Diabetes Prevention Clinical Trials

Briefing Dossier for Qualification Opinion

Submitted to the European Medicines Agency

September 13, 2020

Type 1 Diabetes Consortium (T1DC)

Inish O'Doherty, T1DC Executive Director	Critical Path Institute
Jessica Dunne, Co-Director	Janssen
Marjana Marinac, Co-Director	JDRF
Claudia Richard, Co-Chair	Janssen
Klaus Romero, Co-Chair	Critical Path Institute
Johnna Wesley, Co-Chair	Novo Nordisk

T1D Disease Modeling & Simulation Team Members

Jackson Burton	Critical Path Institute
Sarah David	Critical Path Institute
Patrick Lang	Critical Path Institute
Rhoda Muse	Critical Path Institute
Inish O'Doherty	Critical Path Institute
Jagdeep Podichetty	Critical Path Institute
Klaus Romero	Critical Path Institute
Laura Song	Critical Path Institute

List of Abbreviations

AA	Autoantibodies
ADA	American Diabetes Association
AFT	Accelerated Failure Time
AIC	Akaike's Information Criteria
AUC	Area under the curve
BMI	Body Mass Index
C-Path	Critical Path Institute
COU	Context of Use
Cox PH	Cox Proportional Hazard
DAISY	Diabetes Auto Immunity Study in the Young
DKA	Diabetic ketoacidosis
FDA	US Food and Drug Administration
FDR	First-degree relative
GAD65	Glutamic acid decarboxylase 65 autoantibody
HbA1c	Hemoglobin A1c
HLA	Human leukocyte antigen
IA-2	Insulinoma antigen-2 autoantibody
IAA	Insulin/proinsulin autoantibody
NIDDK	National Institute of Diabetes and Digestive and Kidney Diseases
OGTT	Oral Glucose Tolerance Test
PH	Proportional Hazard
Q-Q	Quantile-Quantile
ROC	Receiver operating characteristic
T1D	Type 1 Diabetes
T1DC	Type 1 Diabetes Consortium
TEDDY	The Environmental Determinants of Diabetes in the Young
TN01	TrialNet Pathway to Prevention Study
VPC	Visual Predictive Check
WHO	World Health Organization
ZnT8	Zinc transporter 8 autoantibody

1 Table of Contents

1 3
1.1 List of Tables 4
1.2 List of Figures 4
2 EXECUTIVE SUMMARY 5
2.1 The Objective(s) of Request 5
2.2 The Need and Impact of Proposed Clinical Novel Methodologies 5
2.3 Characteristics of the Proposed Novel Methodology 7
2.4 Sources of Data and Major Findings 7
2.5 Overall Goal of the Present Submission 8
2.6 Conclusion 8
3 BACKGROUND 8
3.1 Regulatory History 8
3.2 Proposed Context-of-Use Statement 9
4 METHODOLOGY AND RESULTS 9
4.1 Introduction 9
4.2 Context-of-Use 10
4.3 Methods 11
4.3.1 Data 11
4.3.2 Derivation of Analysis Set and External Validation Set 17
4.3.3 Data Analysis 21
4.3.4 Modeling Analysis Methodologies 26
4.3.5 Cox Proportional Hazard Model 28
4.3.6 Parametric Accelerated Failure Time Model 29
4.3.7 Model Performance and Validation 31
4.4 Results 32
4.4.1 Cox Proportional Hazard Model 33
4.4.2 Parametric Accelerated Failure Time Model 36
4.4.3 Model Performance and Validation 42
4.5 Intended Application of Proposed Tool 46
4.5.1 Methodology of Tool 46
4.5.2 Key Deliverables 47
4.6 Conclusions: Methodology and Results 47
5 SUMMARY AND CONCLUSIONS 48
6 QUESTIONS FOR EMA FOLLOWED BY T1DC's POSITION 50
7 REFERENCES 52

1.1 List of Tables

Table 1. Overview TN01, TEDDY, and DAISY 12
Table 2. Patient features at recorded diagnosis times in each study 18
Table 3. Covariates evaluated 20
Table 4. Data summary of covariates and diagnoses by study for analysis set 22
Table 5. T1D diagnoses in the analysis set by autoantibody combination 22
Table 6. HLA Risk Categories 24
Table 7. Diagnosis information for individuals with varying number of islet AAs 26
Table 8. Historical quantitative modeling for predicting T1D diagnosis 28
Table 9. Survival function with various forms of parametric distributions 30
Table 10. Univariate analysis for each covariate using Cox PH Model 33
Table 11. Wilcoxon test between continuous and categorical covariates 34
Table 12. Chi-square test of independence between categorical covariates 35
Table 13. Value of AIC for models fitted with Cox PH 35
Table 14. Selected Cox PH model parameter estimates (model 6) 35
Table 15. Testing proportional hazards assumption using Schoenfeld residuals to test for independence between residuals and survival time 36
Table 16. AIC value for hazard distribution 37
Table 17. Univariate analysis for each covariate using AFT model with Weibull distribution 38
Table 18. Values of AIC for AFT models fitted with a Weibull distribution 39
Table 19. Model 6 (orig_mod) parameter estimates 39
Table 20. Value of AIC for original model (model 6) and other alternative models 40
Table 21. Selected model (alt_mod3) parameter estimates 40
Table 22. C-index values over 6 years for each fold during k-fold cross validation analysis 44
Table 23. C-index values over six years with cross-validation on a pediatric population (age < 12) 45
Table 24. C-index values over six years with DAISY external validation dataset 46

1.2 List of Figures

Figure 1 A - E. Continuous covariates stratified by diagnoses 23
Figure 2. Distribution of time difference between T1D diagnosis and covariates 25
Figure 3. Modeling development workflow 27
Figure 4. Pearson's correlation between continuous covariates 34
Figure 5. Graphical diagnostics with scale Schoenfeld residuals (Beta(t)) against survival time 36
Figure 6. Survival and hazard plots for Weibull Distribution. (Black lines represent the estimated hazard and survival functions from the data, and red solid and dotted lines indicate the mean and 95% confidence intervals of the Weibull model fit.) 37
Figure 7. QQ plots for categorical covariates in the model 41
Figure 8. Evaluation of model performance using time dependent receiver operation characteristic (ROC) analysis 43
Figure 9. VPC-style plots for k-fold cross validation (red shaded region shows the 95% prediction interval and the black shaded region shows the 95% CI for the observed data) 44
Figure 10. VPC-style plot for internal cross validation (CV) using pediatric population (red shaded region shows the 95% prediction interval and the black shaded region shows the 95% CI for the observed data) 45
Figure 11. VPC-style plot for external validation using the DAISY analysis dataset (red shaded region shows the 95% prediction interval and the black shaded region shows the 95% CI for the observed data) 46

2 EXECUTIVE SUMMARY

2.1 The Objective(s) of Request

The objective of this briefing dossier is for the Critical Path Institute's Type 1 Diabetes Consortium (T1DC) to achieve a qualification opinion for a new drug development tool for Type 1 Diabetes (T1D) through EMA's qualification of novel methodologies for medicine drug development. This dossier contains the proposed context-of-use (COU) statement, data source description, modeling analysis methods and results that provide a quantitative basis to support the use of islet autoantibodies (AAs) to enrich subjects for inclusion in T1D prevention trials. An accelerated time failure model will provide the supporting evidence for the use of islet AAs anti-insulin AA (IAA), anti-glutamic acid decarboxylase 65 AA (GAD65), anti-insulinoma antigen-2 AA (IA-2), and zinc transporter 8 AA (ZnT8) as enrichment biomarkers in T1D prevention clinical trials. The presence of different numbers and combinations of islet AAs were analyzed in conjunction with other relevant sources of variability including, demographics, human leukocyte antigen (HLA) haplotype, first-degree relative (FDR), T1D status and blood glucose assessments. The specific sources of variability that were selected include baseline age, sex, blood glucose measurements from the 120-minute timepoints of an Oral Glucose Tolerance Test (OGTT), and hemoglobin A1c (HbA1c) test. The process by which these sources of variability were selected is outlined in this briefing dossier.

2.2 The Need and Impact of Proposed Clinical Novel Methodologies

T1D is a chronic autoimmune disease that results from the destruction of insulin-producing beta cells (β-cells) in the islets of Langerhans of the pancreas. The ability of T1D patients to make insulin is impaired, and consequently patients are unable to regulate their blood glucose levels. T1D affects 3 million people in Europe (International Diabetes Federation 2019). The incidence of T1D is on the rise worldwide, particularly in children. In Europe, incidences vary between 0.2 and 0.5%, with steep rises in the number of children and young people with T1D in lower incidence countries like Hungary and Poland, catching up the high incidence countries, like Scandinavia, where stabilization seems to happen (Barkai et al. 2020; Szalecki et al. 2018; Skrivarhaug et al. 2014).

Insulin replacement therapy remains the cornerstone of treatment for T1D and is used to manage blood glucose levels. Diabetes-associated complications (nephropathy, neuropathy, retinopathy, and cardiomyopathy) arise from poor long-term glycemic control. Tight glucose control through optimal insulin therapy management can reduce the risk of developing diabetes-related complications (Writing Group for the DCCT/EDIC Research Group et al. 2015) but most patients fail to achieve their glycemic target (Foster et al. 2019). Although insulin therapy and blood glucose management provide substantial benefit to patients (Writing Group for the DCCT/EDIC Research Group et al. 2015), this approach does not target the underlying destructive autoimmune processes that drive disease pathogenesis.

The challenges associated with the development of therapies to prevent or delay the onset of T1D are multifactorial and include a lack of qualified biomarkers that identify individuals at risk of developing T1D or can aid in quantifying the risk of conversion to a T1D diagnosis. There have been significant late-stage failures in the development of therapies in new-onset T1Dwhich have been attributed, in part, to two factors. First, there is a high degree of heterogeneity in the patient population and an inability to quantitatively describe the specific sources of variability that may contribute to this heterogeneity. Second, due to the nature of the disease, intervening in new-onset T1D may be too late to significantly delay or halt disease progression and preserve endogenous β-cell function.

Decades of research into the natural history of T1D has demonstrated that subjects at risk for developing T1D can be identified. Subjects that are FDR of T1D patients or express a specific HLA haplotype (HLA-DR3/3, DR4/4, DR3/4, DR3/X [X≠3], DR4/X [X≠4]) have been shown to be at risk for developing T1D. Further, the presence of multiple islet AAs including IAA, GAD65, IA-2, and ZnT8 have been shown to be robust biomarkers capable of predicting a clinical diagnosis of T1D (Hagopian et al. 2011). The summarized main finding of this work is that in individuals at risk of developing T1D (those with specific HLA haplotype or who are a FDR), the presentation of two or more islet AAs (A. G. Ziegler et al. 2013) will eventually lead to the onset of T1D over time and the rate of conversion to a T1D diagnosis is increased with a greater number of islet AAs (Sosenko et al. 2009; Veijola et al. 2016; Xu, Krischer, and Type 1 Diabetes TrialNet Study Group 2016). Multiple natural history studies (Hagopian et al. 2011; Mahon et al. 2009; Skyler et al. 2005; Insel et al. 2015) have been utilized to develop a staging classification guideline, endorsed by JDRF, American Diabetes Association (ADA), and the Endocrine Society, that aids in the description of T1D in its pre-diagnosis stages. This classification is a tool that provides a common language for the T1D community to discuss the stages preceding the clinical diagnosis of T1D and to build awareness among clinicians, as well as patients and their relatives at higher risk of progressing to a clinical disease (Insel et al. 2015). Similar natural history studies are ongoing in the general pediatric population (i.e. subjects without FDR or specific HLA risk haplotype) to assess the prevalence of the islet AAs and their relationship to T1D diagnosis ("ASK Research Program / Autoimmunity Screening for Kids / Denver, CO" n.d.; Kick et al. 2018; Raab et al. 2016; A. Ziegler et al. 2020).

However, translating the findings from these natural history studies to comprehensively inform subject selection in T1D prevention trials is challenging due to the variability of the latency phase, defined as the time from presentation of multiple islet AAs to ultimate diagnosis of T1D. To utilize islet AAs as enrichment biomarkers in drug development, it is necessary to quantitatively determine if the islet AA combinations are statistically significant predictors of T1D diagnosis. Without this understanding, sponsors are unable to design informative clinical trials of appropriate and reasonable size, duration, and cost that will be capable of adequately evaluating potentially transformational therapies.

In order to address this drug development need, the T1DC 1) acquired, remapped, integrated and curated existing patient-level data from observational studies and 2) evaluated the utility of islet AAs, including IAA, GAD65, IA-2, and ZnT8 as biomarkers to enrich subjects for inclusion in T1D prevention trials using a model-based approach. With the patient-level data available to the T1D Consortium's team, this model-based approach considered other sources of variability 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. The output of this model-based approach is a quantitative description of time-varying probability of reaching a diagnosis of T1D that could be used to optimize subject enrichment strategies for T1D prevention trials aiming to delay or prevent T1D.

Based on the methods and results outlined in this briefing dossier, the T1DC is seeking qualification opinion for islet AAs as enrichment biomarkers for T1D prevention clinical trials. This qualification will also help to de-risk drug development and streamline the review of new drug candidates for T1D. The T1DC believes qualification of these biomarkers is particularly important as there are currently many new marketing authorization applications seeking review.

2.3 Characteristics of the Proposed Novel Methodology

Proposed Context-of-Use Statement

In individuals at risk of developing T1D, the islet AAs can be used together with other patient features, as enrichment biomarkers to optimize the selection of individuals for clinical trials of therapies intended to prevent or delay the clinical diagnosis of T1D. The islet AAs proposed include IAA, GAD65, IA-2, and ZnT8. Additional patient features include sex, baseline age, blood glucose measurements from the 120-minute timepoints of OGTT and HbA1c levels.

- **General Area:**

  Enrichment biomarkers for clinical trials focusing on the delay or prevention of the clinical diagnosis of T1D.

- **Target Population for Use of the Biomarkers:**

  Individuals at risk of T1D, defined as being a FDR of a T1D patient, or having a specific HLA subtype of risk (HLA-DR3/3, DR4/4, DR3/4, DR3/X [X≠3], DR4/X [X≠4]). It is intended that positivity for two or more of the islet AAs be determined in this population, to be used as enrichment biomarkers for clinical trials focusing on the delay or prevention of the clinical diagnosis of T1D.

- **Stage of Drug Development for Use:**

  All clinical efficacy evaluation stages of therapeutic interventions focused on the prevention or delay of T1D, including early signs of efficacy, proof-of-concept, dose-ranging, and registration studies.

- **Intended Application:**

  To utilize the islet AAs as enrichment biomarkers as a means of patient selection in clinical trials investigating therapies that are intended to prevent or delay the clinical diagnosis of T1D. These biomarkers, along with additional patient features, such as baseline HbA1c levels and the 120-minute timepoint from an OGTT, can be used as predictors to identify subpopulations at highest risk of a diagnosis of T1D during the course of T1D prevention clinical trials.

2.4 Sources of Data and Major Findings

As of May 2020, the T1DC has obtained three datasets, The Environmental Determinants of Diabetes in the Young (TEDDY), the TrialNet Pathway to Prevention Study (TN01) and the Diabetes Autoimmunity Study in the Young (DAISY). The TEDDY and TN01 were aggregated to support the model-based qualification of islet AAs as enrichment biomarkers. This aggregated dataset was used to construct the statistical analysis plan presented in the T1DC's May 2019 submission for qualification advice. The developed model demonstrates that the islet AAs are statistically significant predictors of the time-varying probability of conversion to a diagnosis of T1D. Further when additional sources of variability, including baseline age, sex, blood glucose measurements from the 120-minute timepoints of OGTT, and HbA1c, are assessed with the islet AAs, it further improves the accuracy of predicting the time-varying probability of conversion to a T1D diagnosis. Since the May 2019 submission, the T1DC has acquired the data from DAISY which was reserved to externally validate the model. In summary, analysis of TN01, TEDDY, and DAISY, constitute data-driven evidence for using the presence of two or more islet AAs and other patient features as enrichment biomarkers for selection of subjects included in T1D prevention studies.

2.5 Overall Goal of the Present Submission

The T1DC presents this briefing dossier to obtain the Agency's Qualification Opinion on the proposed COU for the islet AAs as enrichment biomarkers for T1D prevention trials. The T1DC believe the Qualification Opinion will be critical for the acceleration of the development of drugs that prevent or delay the onset of T1D.

2.6 Conclusion

The developed model was shown to demonstrate that the presence of two or more islet AAs are statistically significant predictors of the time-varying probability of conversion to a diagnosis of T1D. Furthermore, glycemic measurements within this multiple islet AA positive population were shown to further contribute as independent predictors thereby increasing the accuracy of predicting the time-varying probability of conversion to a T1D diagnosis. The T1DC team considers that this model provides the supporting evidence for the application islet AAs as enrichment biomarkers as defined by the context of use statement.

3 BACKGROUND

3.1 Regulatory History

This dossier reports the result of the T1DC and Critical Path Institute's analysis of islet AAs as enrichment biomarkers in clinical trials for the prevention or delay of T1D. Previous interactions between T1DC and EMA's Scientific Advice Working Party (SAWP) are as follows:

- May 2019: T1DC submitted a briefing dossier for qualification advice
- September 2019: Face-to-face discussion meeting between SAWP and T1DC
- October 2019: EMA issued T1DC qualification advice
- March 2020: EMA issued Letter of Support for islet AAs as enrichment biomarkers for T1D prevention studies
- June 2020: T1DC submitted briefing dossier for qualification opinion to SAWP
- July 2020: T1DC received List of Issues from SAWP

This submission now seeks a qualification opinion from EMA regarding the proposed novel methodology (islet AAs as enrichment biomarkers in clinical trials for the prevention or delay of T1D). A full and detailed description of the initial submission for qualification advice can be found in Appendix F. Following submission of this briefing dossier in June 2020, a List of Issues adopted by the SAWP during its 6-9 July 2020 meeting. The SAWP List of Issues included a request for the T1DC to address the first three issues separately in writing, while the remaining issues will be discussed during a discussion meeting with SAWP scheduled for the week of September 28,2020. The first three issues to be addressed in writing pertained to statistical notation of the model, exploration of alternative models by adding baseline age and sex in different combinations to the selected AFT model, and provision of "visual predictive check" style figures. As per the request, the written responses to these three issues are included in a separate document, as well as in Sections 4.3.6 and 4.3.7 of this document. The Methods and Results sections of this briefing dossier were subsequently updated based on feedback from the List of Issues, as were Appendix H and code files within the Modeling Analysis Zip file.

T1DC's previous regulatory history with FDA's Biomarker Qualification Program (BQP) are as follows:

- August 31, 2019: T1DC submitted Letter of Intent to FDA BQP
- March 14, 2019: BQP issues a favorable Determination Letter accepting islet AAs as an enrichment biomarker into the BQP.

In its response letter, FDA suggested categorizing the islet AAs as susceptibility-risk markers for the enrichment of subjects in T1D prevention trials, rather than the T1DC's initial prognostic biomarker categorization. The T1DC is currently preparing a Stage 2 Qualification Plan submission to FDA while it pursues a Qualification Opinion with EMA. A full and detailed description of FDA's comments on the Letter of Intent submission can be found in Appendix G.

3.2 Proposed Context-of-Use Statement

Proposed Context-of-Use Statement

In individuals at risk of developing T1D, the islet AAs can be used together with other patient features, as enrichment biomarkers to optimize the selection of individuals for clinical trials of therapies intended to prevent or delay the clinical diagnosis of T1D. The islet AAs proposed include IAA, GAD65, IA-2, and ZnT8. Additional patient features include sex, baseline age, blood glucose measurements from the 120-minute timepoints of OGTT and HbA1c levels.

- **General Area:**

  Enrichment biomarkers for clinical trials focusing on the delay or prevention of the clinical diagnosis of T1D.

- **Target Population for Use of the Biomarkers:**

  Individuals at risk of T1D, defined as being a FDR of a T1D patient, or having a specific HLA subtype of risk (HLA-DR3/3, DR4/4, DR3/4, DR3/X [X≠3], DR4/X [X≠4]). It is intended that positivity for two or more of the islet AAs be determined in this population, to be used as enrichment biomarkers for clinical trials focusing on the delay or prevention of the clinical diagnosis of T1D.

- **Stage of Drug Development for Use:**

  All clinical efficacy evaluation stages of therapeutic interventions focused on the prevention or delay of T1D, including early signs of efficacy, proof-of-concept, dose-ranging, and registration studies.

- **Intended Application:**

  To utilize the islet AAs as enrichment biomarkers as a means of patient selection in clinical trials investigating therapies that are intended to prevent or delay the clinical diagnosis of T1D. These biomarkers, along with additional patient features, such as baseline HbA1c levels and the 120-minute timepoint from an OGTT, can be used as predictors to identify subpopulations at highest risk of a diagnosis of T1D during the course of T1D prevention clinical trials.

4 METHODOLOGY AND RESULTS

4.1 Introduction

The purpose of this analysis is to develop a time-to-event model that describes the time-varying probability of T1D onset in at-risk subjects. The developed model accounts for several predictors of T1D diagnosis within the defined patient population (individuals that are FDRs of a T1D patient, or who have an HLA haplotype of risk, defined as HLA-DR3/3, DR4/4, DR3/4, DR3/X [X≠3], DR4/X [X≠4]). This model is intended to provide the necessary evidence to support the use of islet AAs to enrich trials with subjects who have a higher likelihood of

reaching T1D diagnosis over time based on the positivity of two or more islet AAs, 120-minute timepoint values from an OGTT, and HbA1c levels, for T1D prevention studies.

The objectives of this analysis are:

- To develop a time-to-event model to predict the time-varying probability of T1D diagnosis in individuals that are FDRs of a T1D patient, or who have an HLA haplotype of risk
- To leverage the model to quantify the effect of various combinations of islet AAs positivity, together with glycemic measures on time-varying probability of T1D diagnosis
- To allow the determination of optimal enrichment strategies in clinical trials intended to prevent or delay the onset of T1D
- To develop an open-source accelerated failure time (AFT) survival model based in the R programming language to allow for the use of the developed model for clinical trial enrichment strategies

4.2 Context-of-Use

Proposed Context-of-Use Statement

In individuals at risk of developing T1D, the islet AAs can be used together with other patient features, as enrichment biomarkers to optimize the selection of individuals for clinical trials of therapies intended to prevent or delay the clinical diagnosis of T1D. The islet AAs proposed include IAA, GAD65, IA-2, and ZnT8. Additional patient features include sex, baseline age, blood glucose measurements from the 120-minute timepoints of OGTT, and HbA1c levels.

- **General Area:**

  Enrichment biomarkers for clinical trials focusing on the delay or prevention of the clinical diagnosis of T1D.

- **Target Population for Use of the Biomarkers:**

  Individuals at risk of T1D, defined as being a FDR of a T1D patient, or having a specific HLA subtype of risk (HLA-DR3/3, DR4/4, DR3/4, DR3/X [X≠3], DR4/X [X≠4]). It is intended that positivity for two or more of the islet AAs be determined in this population, to be used as enrichment biomarkers for clinical trials focusing on the delay or prevention of the clinical diagnosis of T1D.

- **Stage of Drug Development for Use:**

  All clinical efficacy evaluation stages of therapeutic interventions focused on the prevention or delay of T1D, including early signs of efficacy, proof-of-concept, dose-ranging, and registration studies.

- **Intended Application:**

  To utilize the islet AAs as enrichment biomarkers as a means of patient selection in clinical trials investigating therapies that are intended to prevent or delay the clinical diagnosis of T1D. These biomarkers, along with additional patient features, such as baseline HbA1c levels and the 120-minute timepoint from an OGTT, can be used as predictors to identify subpopulations at highest risk of a diagnosis of T1D during the course of T1D prevention clinical trials.

4.3 Methods

4.3.1 Data

The T1DC has obtained three datasets, the TN01, TEDDY, and DAISY studies, to support the qualification of islet AAs as enrichment biomarkers. TEDDY and TN01 were aggregated and used for model development and internal cross-validation. Based on the results of the modeling analysis, these two datasets were considered sufficient for this purpose. Data from the DAISY study was acquired and used to perform external validation on the selected model. All studies are observational but certain features in their designs differ, including inclusion criteria and scheduled frequency of follow-up. A summary of the three studies can be found in Table 1.

For the three studies, serum was collected from participants and was analyzed using radio-ligand binding assays to determine the binary, qualitative output of seropositivity or negativity for an individual autoantibody. Assays were run in centralized labs that used either a series of positive controls derived from subjects with recently diagnosed T1D and negative controls based on healthy subjects or a series of reference standards provided by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) to establish cut points for determining if an individual was either seropositive or negative for a specific autoantibody. Full details for each assay can be found in Appendices A, B, and Addendum 1.

These longitudinal studies monitored large numbers of subjects over many years for the presentation of the islet AAs, with the predominant result during this monitoring phase being seronegativity for individual islet AAs. To avoid the prohibitive cost of reassessing all seronegative results, the studies focused their reassessments on those subjects who demonstrated seropositivity for a particular islet AA.

4.3.1.1 Studies

Type 1 Diabetes TrialNet is an international consortium of clinical research centers that participate in studies designed to further the goals of prevention or delay of T1D. TrialNet began with the Pathway to Prevention Study (TN01) (previously called the natural history study) and has progressed into a clinical trial consortium that executes interventional trials in both T1D prevention and in newly diagnosed T1D patients. The overall objective of TN01 is to perform baseline and repeat assessments over time of the metabolic and immunologic status of individuals at risk for T1D to: (a) characterize their risk for developing T1D and identify subjects eligible for prevention trials, (b) describe the pathogenic evolution of T1D, and (c) increase the understanding of the pathogenic factors involved in the development of T1D. The primary outcome of TN01 is the development of diabetes as defined by the ADA based on glucose testing, or the presence of symptoms and unequivocal hyperglycemia.

Participants for TN01 were selected by the presence of a FDR with T1D, as this has been shown to be a risk factor for the development of T1D. The criteria included (1) FDRs (age 1–45 years) of T1D probands or (2) second- and third-degree relatives (age 1–20 years) of T1D probands (i.e., nieces, nephews, aunts, uncles, grandchildren, cousins, half-siblings). Based on these criteria, 211,230 subjects with positive FDRs were screened for the presence of islet AAs, as of November 2018. Between 2004-2009 subjects with the presence of one islet AA were considered eligible for follow-up. In 2009 the eligibility criteria for follow-up changed to the presence of two islet AAs. Once subjects were selected for follow-up and opted in, they were monitored at six-monthly visits using OGTT, detection of islet AAs, and measurement of HbA1c levels.

Table 1. Overview TN01, TEDDY, and DAISY

	TN01	TEDDY	DAISY
Type of study:	Observational	Observational	Observational
Years running:	2004-Present	2004-Present	1993-Present
Enrollment design:	Ongoing screening and active enrollment	Screening complete and fixed prospective cohort	Screening complete and fixed prospective cohort
Enrollment criteria:	Ages 1-45 must have FDR with T1D*, ages 1-20 must have extended family member** with T1D	Newborns (< 4 months old) with high-risk HLA*** or FDR	Newborns with high-risk HLA or FDR Sibling/offspring of individual with T1D, initial visit <7yo
Number of subjects:	209,394 initial screening 4,524 being followed (December 2018)	361,518 initial screening 8,667 in initial prospective cohort	31,881 initial newborn screening 2,547 in prospective cohort.
Primary Study Outcome:	T1D diagnosis	Appearance of one or more islet cell autoantibodies	T1D diagnosis
Secondary Study Outcome:	Metabolic and autoantibody measurements	T1D diagnosis	Detection of islet autoantibodies
Average age at entry:	19.1 years (<3 months to >49 years)	3 months	Average age at entry for newborn screened: 1.0 yr Average age at entry for sib/offspring cohort: 2.31 yr
Number of subjects who tested positive for 1 islet AA at or after screening:	13,058†	794	364
Number of subjects who tested positive for 2 islet AAs at or after screening:	4,550	535	136

* FDR is defined as a child, parent, or sibling.
** Extended family member is defined as a cousin, niece, nephew, aunt, uncle, grandparent, or half-sibling.
*** High risk HLA is defined as having an HLA genotype that is associated with higher incidences of HLA. In the TEDDY study these were HLA-DR3/3, DR4/4, DR3/4, DR3/X [X≠3], DR4/X [X≠4]
† Between 2004-2009 individuals with one islet AA were followed with six-monthly assessments. After 2009 this changed, and subjects required two or more islet AAs to be enrolled in the follow-up cohort

TN01 is still enrolling new subjects and following current subjects, however the TN01 data provided in this submission are locked as of December 2018. TN01 is providing TrialNet with an active patient ready cohort and collaborative clinical trial network to evaluate novel therapies for immune modulation and/or enhancement of β-cell proliferation and regeneration. These interventional studies (Battaglia et al. 2017) are investigating therapies designed to prevent T1D or therapies to preserve β-cell function in individuals with newly diagnosed T1D who have residual β-cell function. A full copy of the TN01 protocol can be found in the TN01 Manual of Operations (Appendix C).

Inclusion criteria: TN01 is divided into three phases: Screening of subjects with positive FDRs (Phase 1), Baseline Risk Assessment (Phase 2) and Follow-up Risk Assessments (Phase 3). Phase 1 (211,230 individuals) involves overall screening and biochemical measurements of islet AAs to determine eligibility for the Phase 2 risk assessment (6,297 individuals). Once an individual is found to have two or more islet AAs (between 2004-2009 this was one more islet AA) they are offered to participate in Phase 3 where they are monitored at six-month intervals (4,524 individuals).

Baseline assessment: Phase 1 and 2 will establish a baseline assessment which will include a categorization of FDR, an OGTT, measurement of HbA1c, testing for islet AAs, and HLA typing.

Follow up assessments: Subjects will be seen at six-month intervals for the duration of the study for Phase 3 follow-up risk assessments. At each visit, tests will include OGTT, collection of blood for islet AA testing and measurement of HbA1c levels.

Determination of islet autoantibody positivity: A full and detailed description of the performance characteristics and cutoffs used to assess seropositivity in each islet AA assay can be found in Appendix A and Addendum 1. Importantly, for any serum sample called positive, a follow-up confirmatory test is performed. For the TN01 study, a Laboratory Monitoring Committee (LMC) is responsible for monitoring assay accuracy and consistency of each TrialNet participating laboratory. Each laboratory participating in TN01 is reviewed by a LMC every six months; see Appendix A for most recent quality control reports for the Denver lab (dated July/August 2018). A brief summary of the quality control procedures for TN01 follows:

1. All samples are always run in duplicate.
2. Three internal standard control samples (one high positive control, one low positive control, and one negative control) are included in each assay. A set of NIDDK standard samples for a standard curve are also included in each assay for harmonized GAD65 and IA-2 assay protocols. All internal standard control samples and a NIDDK standard curve sample set are included in every four plates if more than four plates are being tested (Recommendation by NIDDK islet Autoantibody Harmonization Committee).
3. Every positive result is confirmed by re-run in the different assay on a different day, again in duplicate. If the 2nd run comes out negative, which disagrees with the result of 1st run, a 3rd run will be necessary. The average value of two runs, which agree (both positive or both negative), is the final result for data entry and report. Still, all results from these different assays are available in a database upon request.
4. The lab does not re-test negative samples as it is rare to see negative results re-test as positive (i.e. the frequency is extremely low and since most screening samples are negative so it would be very costly to re-test.

Endpoint in protocol: T1D diagnosis as classified by the presence of unequivocal hyperglycemia including acute metabolic decompensation (diabetic ketoacidosis), or

The following criteria must be met on two occasions:

1. Signs and symptoms of diabetes plus casual plasma glucose concentration of ≥200 mg/dL (11.1 mM). Casual is defined as any time of day without regard to time since last meal. The classic symptoms of diabetes include polyuria, polydipsia, and unexplained weight loss, or:
2. Fasting Plasma Glucose (FPG) ≥126 mg/dL (7.0 mM). Fasting is defined as no caloric intake for at least 8 hours, or
3. 2-hour Plasma Glucose (PG) ≥200 mg/dL (11.1 mM) during an OGTT.

The Environmental Determinants of Type 1 Diabetes in the Young (TEDDY) is a prospective cohort study of 8667 children at high genetic risk for T1D, which seeks to identify environmental causes of T1D (Hagopian et al. 2006)). There are six clinical research centers: three in the U.S. (Colorado, Georgia/Florida, Washington), and three in Europe (Finland, Germany, and Sweden). Children were screened and recruited during infancy based on high-risk HLA genotypes, with separate inclusion criteria for the general population (GP) children or FDR as described (Hagopian et al. 2011). Data for the high-risk HLA genotypes and FDRs, but not the GP, will be included as part of this submission.

TEDDY is longitudinally prospective study assessing a broad spectrum of environmental factors that may contribute to the stimulus or stimuli that are involved in the immune initiation of T1D. An assessment of these environmental factors that will not be part of this submission, include identification of infectious agents, dietary factors, or other environmental agents, including psychosocial factors. Of participants, 89% had no family history of T1D. Participants are monitored prospectively with study visits every three months for the first four years, and every six months thereafter to age 18. All children who are persistently positive for any islet AA are monitored every three months until the age of 15 years or diagnosis of T1D. As of November 2018, 9.1% of the participants had developed at least one islet AA; 3.8% had developed T1D and thus reached study endpoint. Of the original cohort who have not reached the study endpoint, 68% are still participating in follow-up. TEDDY data provided in this submission are locked as of June 2018. A full copy of the TEDDY protocol can be found in the TEDDY Manual of Operations (Appendix D).

Inclusion criteria: 8668 new-born subjects (0-3 months) were screened for high-risk HLA genotypes. High-risk HLA types in TEDDY are classified as HLA-DR3/3, DR4/4, DR3/4, DR3/X [X≠3], DR4/X [X≠4]. A detailed definition of the high-risk HLA typing assignments can be found in Hagopian et al. 2011.

Baseline assessment: Once a subject has been deemed HLA eligible via baseline assessment, and subjects' guardians have consented, the subject is entered into the follow-up cohort.

Follow-up assessments: Subjects are assessed every three months for the first four years of life with a variety of physiological lab tests to determine the environmental factors effecting the immune-initiation of T1D. For this submission, only the positivity for islet AAs and blood glucose measures were considered. If a subject becomes positive for an islet AA they will continue on a three-month visit schedule until 15 years of age or the diagnosis of T1D. If a subject is negative for islet AAs after the first four years, the subject will shift to a six-month assessment schedule but will shift back to a quarterly assessment if they become islet AA positive. OGTTs will only be done on every subject that is >3 years of age with two islet AAs.

Determination of islet autoantibody positivity: A full and detailed description of the performance characteristics and cutoffs used to assess seropositivity in each islet AA assay can be found in Appendix A and Addendum 1. Importantly, for any serum sample determined to be positive, a follow-up confirmatory test is performed. In TEDDY, two centralized labs are utilized to confirm samples that are positive for an individual islet AA. Samples that are tested as positive in Denver are sent to Bristol for confirmation, and samples tested as positive in Bristol are sent to Denver for confirmation. According to the study protocol, each laboratory should repeat all positive samples internally before reporting positive or negative, and will measure twice, if specific islet AA positives are confirmed, and up to three times if there is a discrepancy between the initial positive result and the second determination (2/3 internal lab reported as positive, with mean of consensus positives or negatives reported in World Health Organization (WHO) units. Results are then sent electronically to the Data Coordinating Center. The Data Coordinating Center then sends the NIDDK repository the ID for all positive samples and a subset of negative samples (5%), and the repository sends the second aliquot of serum to the alternate reference laboratory for confirmation of the positive result.

Endpoint(s) in protocol: The first endpoint in TEDDY is the development of an islet AA.

The final endpoint in TEDDY is T1D diagnosis, classified by the following ADA criteria, which must be met on two occasions (unless criteria 4 is present):

1. Casual (any time of day without regard to time since last meal) plasma glucose >= 200 mg/dL, if accompanied by unequivocal symptoms (i.e. polyuria, polydipsia, polyphagia, and/or weight loss.), or
2. Fasting (no caloric intake for at least 8 hours) plasma glucose >= 126 mg/dL, or
3. 2-hour plasma glucose >=200 mg./dL OGTT. Glucose dose is determinant on body weight to a maximum of 75 grams, or
4. Unequivocal hyperglycemia with acute metabolic decompensation (i.e. ketoacidosis)

Unless criterion 4 is present or the fasting glucose is >=250 mg/dL (at the bedside or in the local laboratory on the day of testing), it is preferred that at least one of the two testing occasions involve an OGTT. If the first criterion met is #3 (i.e. by the 2-hour OGTT value) the OGTT should be repeated within 60 days. It is essential that every effort be made to obtain the necessary tests to establish the diagnosis of diabetes. Subjects will be instructed to eat a balanced diet in the days leading up to the OGTT.

Diabetes Autoimmunity Study in the Young (DAISY) is a prospective cohort study of 2547 children who are at increased genetic risk for developing T1D. DAISY seeks to understand the environmental triggers for islet autoimmunity and progression to T1D. Children were screened and recruited in two groups (1) during infancy based on high-risk HLA genotypes or (2) during early childhood based on first-degree relative (FDR) status as described (Rewers, Norris, et al. 1996; Rewers, Bugawan, et al. 1996).

Children in DAISY were monitored longitudinally for over 20 years, assessing a variety of environmental factors that may be involved in the development of islet autoimmunity. These included assessment of prenatal exposures, birth events, growth and puberty, dietary assessment, smoke exposure, daycare exposure, physical activity assessment, and biological samples for assessment of biomarkers and infectious agents (blood, urine, saliva, throat and rectal swabs).

Participants were assessed at 9, 15 and 24 months of age and then annually thereafter. Those who developed islet autoimmunity were monitored every 6 months. Participants who were positive for more than one islet autoantibody were requested to follow up every 3 months until diagnosis of T1D. As of January 2020, 9.2% of the participants had developed at least

one islet autoantibody and 4.2% had developed T1D. Of the original cohort, 42% were still engaged in follow-up. DAISY data provided in this submission are locked as of June 30, 2017.

More information related to the assays can be referenced in Addendum 1 to Appendix A & B. The protocol for DAISY can be found in Appendix E.

Inclusion criteria:

1) General Population Cohort: Children born November, 1993 through August, 2004 at St. Joseph's Hospital, Denver were recruited in the hospital within days of birth. Cord blood was screened for high-risk HLA genotypes. All children with high- or moderate-risk HLA genoptypes: DR3/4,DQB1*0302, DR3/3, DR4/4 DQB1*0302 or DR4,DQB1*0302/X were invited to participate in follow-up (X≠DR3 or DR4,DQB1*0302). The study enrolled 1555 eligible children, of whom 131 had a first-degree relative with T1D. Included in this total were also 176 children with the DR3/x genotype that is neutral for T1D risk, but a susceptibility genotype for celiac disease. These children could have their first DAISY visit between 9-24 mo of age, while the remainder of this cohort had the initial visit at 9 mo of age.

2) Family History Cohort (FDR): Starting February 11, 1994, young siblings and offspring of a person with T1D were recruited from: i) families of children diagnosed with T1D below age 18, in Colorado, between 1978 and 1991 (Colorado IDDM Registry); ii) families of children with T1D seen in the Barbara Davis Center or The Children's Hospital Colorado after 1991; and iii) media publicity. The FDR (n=995) were invited to participate regardless of their HLA genotype. Most of the FDRs had their initial DAISY visit during the initial 12 months of age; however, the initial visit could be as late as up to 4 y and, in 1993-1995, as late as up to 6.9 y of age.

Exclusion criteria: All: (1) severe co-existent condition, (2) parents both non-English speaking (3) refused consent for long-term storage of data and specimen. General Population Newborns: No cord blood available.

Baseline assessment: Once a subject has been deemed HLA eligible via baseline assessment, and subjects' guardians have consented, the subject is entered into the follow-up cohort.

Follow-up assessments: Participants were assessed at 9, 15 and 24 months of age and then annually thereafter. Participants were assessed for a variety of environmental factors that may be involved in the development of islet autoimmunity. For this submission, only the longitudinal islet AA and blood glucose measures will be considered. If a subject becomes positive for an islet AA they will continue on a 6 month visit schedule until the diagnosis of T1D. Participants who were positive for more than one islet autoantibody were requested to follow up every 3 months until diagnosis of T1D. All islet autoantibody positive participants had HbA1c at each visit. Any participant with 2 or more islet autoantibodies 3 years or older was offered an OGTT every 6 months.

Determination of islet autoantibody positivity: Islet-autoimmunity was assessed from serum sample collected at each clinic visit for radio-binding assay for GAD65, IAA, IA-2 at every visit. Beginning in 2012, participants positive for any other islet antibody or who developed T1D were also tested for ZnT8. Additionally, the last sample collected for all 2547 participants was tested for ZnT8 and if positive, all previous samples were tested to determine age of seroconversion. DAISY participants were tested by radioimmunoassay for GAD65, IAA and IA-2. All available samples from children who were ever positive for any of the above autoantibodies or who developed T1D were tested for ZnT8 as previously described (Wenzlau

et al. 2007). Additionally, the last sample collected for all 2547 participants was tested for ZnT8, and if positive, all previous samples were tested to determine the age of seroconversion. Three children who were positive only for ZnT8 were identified in this way. For any serum sample determined to be positive, a follow-up confirmatory test is performed. Additionally, 10% of negative samples are retested.

Endpoint(s) in protocol: The first endpoint in DAISY is development of an islet AA.

The final endpoint in DAISY is T1D diagnosis, classified by the following ADA criteria that must be met on two occasions (unless criteria 4 is present):

1. Casual (any time of day without regard to time since last meal) plasma glucose >= 200 mg/dL, if accompanied by unequivocal symptoms (i.e. polyuria, polydipsia, polyphagia, and/or weight loss.), or
2. Fasting (no caloric intake for at least 8 hours) plasma glucose >= 126 mg/dL, or
3. 2-hour plasma glucose >=200 mg./dL OGTT. Glucose dose is determinant on body weight to a maximum of 75 grams, or
4. HbA1c ≥ 6.5%
5. Unequivocal symptoms of hyperglycemia

Participants who were diagnosed by a non-study physician outside of the study are included with date of diagnosis and clinical data when available.

4.3.2 Derivation of Analysis Set and External Validation Set

The studies TN01, TEDDY, and DAISY are observational studies focused on monitoring subjects at risk for developing T1D. To perform an analysis using data from these studies, a subset of common variables from all possible variables in each dataset was constructed. The subset of individuals with the common variables, termed the analysis set, were collectively used to inform the modeling analysis for prediction of T1D diagnosis. Based on prior knowledge, subject features relevant prior to T1D diagnosis were selected as part of the analysis set. The list of patient features included in the analysis set were:

- Presence of islet AAs (IAA, GAD65, IA-2, and ZnT8) measured as a binary variable of either seropositivity or seronegativity
- Blood glucose measurements from the 0 and 120-minute timepoints of OGTT tests
- HbA1c measurements
- Demographic information (sex, baseline age, FDR status)
- HLA subtype

In the TN01, TEDDY, and DAISY protocols, the diagnosis of T1D was a study endpoint. The diagnostic criteria pre-specified for each study differed slightly, but each were based on the ADA criteria. Detailed descriptions of the pre-specified diagnosis criteria for each study can be found in Section 4.3.1.1 and in the full TN01, TEDDY, and DAISY protocols (Appendices C, D, and E, respectively). Efforts were then taken to identify and categorize the diagnosis data for the individuals in the analysis set and validation set, 570 diagnoses in 2,061 individuals (Table 2). Table 2 shows a hierarchical breakdown of diagnosis data for subjects present in the analysis set. Data from the three studies showed records of T1D diagnoses that did not fulfill the diagnosis criteria outlined in each study (see Table 2 C - H). Although the T1D diagnoses in categories C-H in Table 2 are not ADA aligned, they were confirmed by a clinician investigator in the TN01, TEDDY, and DAISY studies and were included in the modeling analysis.

Table 2. Patient features at recorded diagnosis times in each study

		TN01		TEDDY		DAISY		Total	
		n	**%**	**N**	**%**	**n**	**%**	**n**	**%**
Individuals with a T1D diagnosis in analysis set:		**398**	**100**	**153**	**100**	**17**	**100**	**570**	**100**
A.	2 positive OGTT*	98	24.6	37	24.2	0	0	135	23.7
B.	Ketoacidosis/Hospitalized at diagnosis	115	28.9	73	47.7	0	0	188	33.0
C.	1 positive OGTT + other positive glucose test****	12	3	16	10.5	2	15.8	31	5.4
D.	1 positive Fasting** + positive HbA1c***	20	5	2	1.3	0	0	22	3.9
E.	Positive HbA1c	48	12.1	5	3.3	4	21.1	57	10.0
F.	Any positive glucose test	71	17.8	5	3.3	3	21.1	80	14.0
G.	Insulin prescribed	18	4.5	0	0	0	0	18	3.2
H.	Only clinician confirmed	0	0	0	0	3	15.8	3	0.5
I.	Diagnosis after 6 years – considered censored	15	3.8	15	9.8	5	26.3	35	6.1
Individuals without a T1D diagnosis in the analysis set:		**1271**	**NA**	**200**	**NA**	**17**	**NA**	**1491**	**NA**
Total number of individuals in analysis set:		**1669**	**NA**	**353**	**NA**	**36**	**NA**	**2061**	**NA**

* Positive OGTT is defined by the ADA as >= 200 mg/dL at 2 hours after oral glucose dose delivery or >=126 mg/dL at the time of oral glucose dose delivery.

** Positive Fasting Glucose is defined as >= 126 mg/dL after 8 hours of fasting.

*** Positive HbA1c is defined as >= 6.5%.

**** Other positive glucose test may include any of the above as well as Random Glucose, defined as >= 200 mg/dL without previous fasting.

4.3.2.1 Definition of Baseline Used for Analysis Set

Only baseline information was used for the modeling analysis. Details regarding consideration of longitudinal analyses are located in Section 4.3.3.3.

The precise definition of baseline used for the analysis set is the first record, i.e. timepoint, for each individual in which the following criteria are satisfied:

- Presence of any two or more islet AAs
- Complete, i.e. non-missing information for OGTT (0 and 120-minute timepoints, HbA1c measurements, age, sex

As part of this derived baseline definition, all variables dependent on the visit day were adjusted. This involved adjusting the age at study entry, diagnosis timing, and last-recorded visit to be based on the newly defined baseline visit day, now labeled as day zero.

The rationale for selecting individuals with only two or more islet AAs is based on the utility of the biomarker for drug development. Individuals with one or fewer islet AAs at baseline have significantly longer expected times to T1D diagnosis. Evidence for this is supported by examining the risk of T1D diagnosis stratified by using only the number of islet AAs present at the first patient record, including zero (see Appendix H Figure 1). For completeness, a supplementary analysis was conducted to assess the feasibility of modeling individuals with one islet AA (see Section 4.3.3.4).

4.3.2.2 Islet AAs as Binary Predictors of T1D Diagnosis

In all three studies, the raw islet AA levels were available. However, as discussed in Section 4.3.1 and Appendices A, B, and Addendum 1, the assays for IAA, GAD65, IA-2, and ZnT8 are only being used to establish the presence or absence for each islet AA, not the continuous quantitative value, as this binary assessment is in line with the fit-for-purpose application of these assays.

To use islet AAs as binary predictors in the model, the islet AAs were represented using dummy variables. Using the requirement that two or more islet AAs must be present, an individual in the analysis subset has exactly one of eleven possible combinations of the four different islet AAs. The binary absence or presence of the islet AAs is then interpreted as one covariate with 11 mutually exclusive levels in which each level is individually assessed for its risk of T1D prediction.

4.3.2.3 Baseline Covariates

Using the definition of the derived baseline as described in the previous section, the baseline covariates evaluated and tested as predictors in the time-to-event model are listed in the following table (Table 3) in terms of their notation and numerical definition.

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.

The final covariates included in the model are listed in Section 4.4.2 after being assessed for their predictive power, and potential collinearity and associations with other covariates.

Table 3. Covariates evaluated

Notation	Description of covariate at derived baseline	Type
$X_{\text{GAD65\_IAA}}$	Positivity for GAD65, IAA	Binary
$X_{\text{GAD65\_IA-2}}$	Positivity for GAD65, IA-2	Binary
$X_{\text{GAD65\_ZnT8}}$	Positivity for GAD65, ZnT8	Binary
$X_{\text{IA-2\_IAA}}$	Positivity for IA-2, IAA	Binary
$X_{\text{IA-2\_ZnT8}}$	Positivity for IA-2, ZnT8	Binary
$X_{\text{IAA\_ZnT8}}$	Positivity for IAA, ZnT8	Binary
$X_{\text{GAD65\_IAA\_ZnT8}}$	Positivity for GAD65, IAA, ZnT8	Binary
$X_{\text{GAD65\_IAA\_IA-2}}$	Positivity for GAD65, IAA, IA-2	Binary
$X_{\text{GAD65\_IA-2\_ZnT8}}$	Positivity for GAD65, IA-2, ZnT8	Binary
$X_{\text{IA-2\_IAA\_ZNT8}}$	Positivity for IA-2, IAA, ZnT8	Binary
$X_{\text{GAD65\_IA-2\_IAA\_ZNT8}}$	Positivity for GAD65, IA-2, IAA, ZnT8	Binary
X_{STUDY}	Flag for being in TN01 or TEDDY	Binary
X_{HR_HLA}	Flag for high risk HLA subtype*	Binary
X_{FDR}	Flag for first-degree relative with T1D **	Binary
X_{SEX}	Male or female	Binary
X_{bAGE_s}	Age	Continuous
X_{BMI_s}	Body mass index	Continuous
X_{HbA1c_s}	HbA1c test result (%)	Continuous
$X_{Log\_GLU0\_s}$	Log transformed and standardized and 0-minute results from OGTT	Continuous
$X_{Log\_GLU120\_s}$	Log transformed and standardized and 120-minute results from OGTT	Continuous

* High-risk HLA is defined in Section 4.3.3.2

** In TN01, the actual FDR was listed, and required a derivation into a binary outcome for the FDR status.

4.3.2.4 Timing of Diagnosis and Removing likely T1D Subjects

A variable was derived, denoted T_event, defined as either the time at which a T1D diagnosis occurred for individuals who had a recorded diagnosis, or the last recorded visit day for individuals with no recorded diagnosis time. For individuals with no recorded diagnosis, T_event is considered the right-censored time, since the event of T1D diagnosis is unobserved.

The analysis set based on the derived baseline includes 168 individuals with OGTT and/or HbA1c values that satisfy the ADA criteria for likely T1D (Fasting Blood Glucose >= 126 mg/dL, Stimulated Blood Glucose >= 200 mg/dL, HbA1c >= 6.5%), but were not assigned a diagnosis at that time. To avoid bias in the modeling analysis, these individuals were counted as false negative and removed from the data representing a reduction in total diagnoses of

146. This censoring resulted in a final total of 551 diagnosis and 2,022 subjects in the analysis set from TN01 and TEDDY and is reflected in Table 2.

4.3.2.5 Missing Data, Imputations, and Right Censoring

The definition of the derived baseline described in Section 4.3.2.1 necessarily excludes all individuals with missing information for OGTT (0 and 120-minute timepoints), HbA1c, sex and age. Therefore, no imputations were required for these variables. In the case of FDR status and HLA subtype, entry criteria differed between TEDDY and TN01 regarding these variables as stated in Section 4.3.1 resulting in missing information in both variables. Additionally, BMI had significant missing information (Section 4.3.3.1). Each of these covariates were nonetheless evaluated for their predictive power, but were not significant predictors and therefore dropped (Section 4.4.1.1). Therefore, in the final modeling analysis, no imputations were performed.

Of the total 551 diagnoses in the analysis set, 30 diagnoses were reached after 6 years. For the parametric modeling approach used in the modeling analysis, an estimate of the baseline hazard function is needed, which is sensitive to sparse information at later diagnosis times. Because of this, diagnosis timing was censored at six years, so that any individual with a diagnosis after six years is considered right-censored with no diagnosis. This approach helps numerically stabilize the estimate of the hazard function and improves model fitting. The approach is further supported due to the lack of importance of accurately predicting diagnoses several years out as such scenarios are not likely for T1D trials of reasonable duration.

4.3.3 Data Analysis

Comprehensive tabulation and visualization of the data contained in the analysis set were performed for reference. Data summaries for the covariates and the diagnosis information for both TN-01 and TEDDY are shown in Table 4. The distribution of diagnosis by islet AA combination are shown in Table 5, and the distribution of diagnosis by each of the continuous covariates are shown in Figure 1. Survival plots were created to display the time-varying incidence of T1D stratified by the covariates in Table 3, shown Appendix H Figure 2-21. These figures show the distribution of right-censored events and the number of individuals at risk by year, which is important for understanding potential informative censoring.

Additional data analysis tasks were done to support the current modeling or to assess if other modeling analyses could be carried out. These included 1) examining the longitudinal information of the islet AAs and the glycemic markers to assess the use of time-varying covariates (Section 4.3.3.2); 2) harmonizing the HLA data across both studies to ensure interoperability; 3) examining the data in individuals with one islet AA to identify potential predictors to T1D diagnosis in shorter durations (<3 years), and Study specific data summaries (Section 4.3.3.3).

4.3.3.1 Data Summaries of TN-01 and TEDDY Analysis Set

In the analysis set, a total of 2,022 subjects were included with complete information for islet AA positivity, age, sex, HbA1c, and 0 and 120- minute timepoints of OGTT. The values of the covariates of interest are shown by study in Table 4. The number of individuals by study for each islet AA combination and the number of diagnoses for that combination are shown in Table 5. Information on FDR, HLA risk group, and BMI contained missing information and is summarized in Table 4. The distributions of continuous covariates (BMI, 0 and 120-minute timepoint OGTT, HbA1c, age) are summarized in Figure 1. The distributions of continuous covariates (BMI, 0 and 120-minute timepoint OGTT, HbA1c, age) based on AA combinations are available in Appendix H figure 22-26.

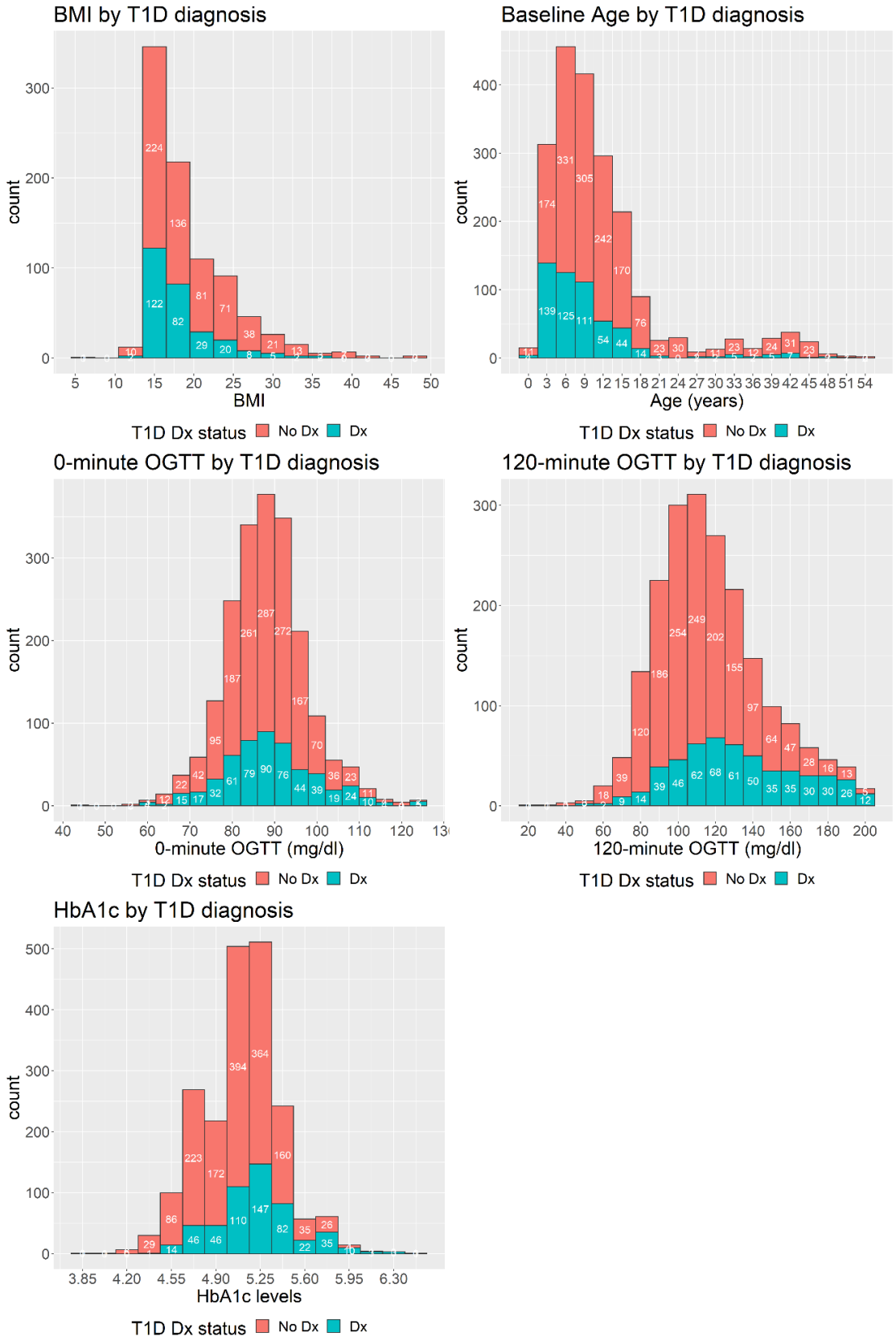
Table 4. Data summary of covariates and diagnoses by study for analysis set

Study	TN01		TEDDY	
	Value	% Missingness	Value	% Missingness
Subjects	1669	-	353	-
Age at Derived Baseline (sd)	13.0 years (10.0)	0	5.7 years (2.5)	0
Sex (% Female)	45.5%	0	41.6%	0.06
Number of Islet AA measurements	1669	0	353	0
Has FDR %	1519	9%	65	0
Mean 0 Min OGTT in mg/dL (sd)	88.9 (9.7)	0	87.0 (8.9)	0
Mean 120 Min OGTT in mg/dL (sd)	120.3 (29.6)	0	108.1 (24.0)	0
HbA1C % (sd)	5.1 (0.3)	0	5.2 (0.2)	0
Number of HLA Measurements	1622	2.8	351	0.6
Mean BMI	21.2 (8.5)	67.6%	16.5 (2.4)	3.1%
Diagnoses	383	NA	138	NA

Table 5. T1D diagnoses in the analysis set by autoantibody combination

	TEDDY			TN01		
Islet AA combination	Subjects	Diagnoses	% Conversion	Subjects	Diagnoses	% Conversion
GAD65_IA-2	34	15	44%	150	35	23%
GAD65_IA-2_IAA	28	13	46%	64	16	25%
GAD65_IA-2_IAA_ZnT8	74	39	53%	280	83	30%
GAD65_IA-2_ZnT8	24	12	50%	315	85	27%
GAD65_IAA	74	15	20%	290	37	13%
GAD65_IAA_ZnT8	26	9	35%	164	28	17%
GAD65_ZnT8	41	3	7%	233	36	15%
IA-2_IAA	10	6	60%	16	4	25%
IA-2_IAA_ZnT8	24	18	75%	51	20	39%
IA-2_ZnT8	12	5	42%	71	32	45%
IAA_ZnT8	6	3	50%	35	7	20%

Figure 1 A - E. Continuous covariates stratified by diagnoses



4.3.3.2 HLA Harmonization

In TEDDY, the HLA subtype was reported as one of the following coded values: HLA-DR3/3, DR4/4, DR3/4, DR3/X [X≠3], DR4/X [X≠4]. These five HLA haplotypes indicate high-risk categories that were used as enrollment criteria for the TEDDY prospective cohort (Hagopian et al. 2011). The TN01 enrollment criteria were based on the presence of a FDR with T1D. Individuals that entered TN01 had the option of being assessed for their HLA status, but not all participants were evaluated. For TN01 individuals in the analysis set that have a determined HLA status, the corresponding coded value of HLA risk used in TEDDY was derived. Any remaining HLA haplotypes in TN01 that did not correspond to the coded values in TEDDY were used to define an alternate risk category for the definition of the covariate assessed in the modeling analysis (Table 6).

Table 6. HLA Risk Categories

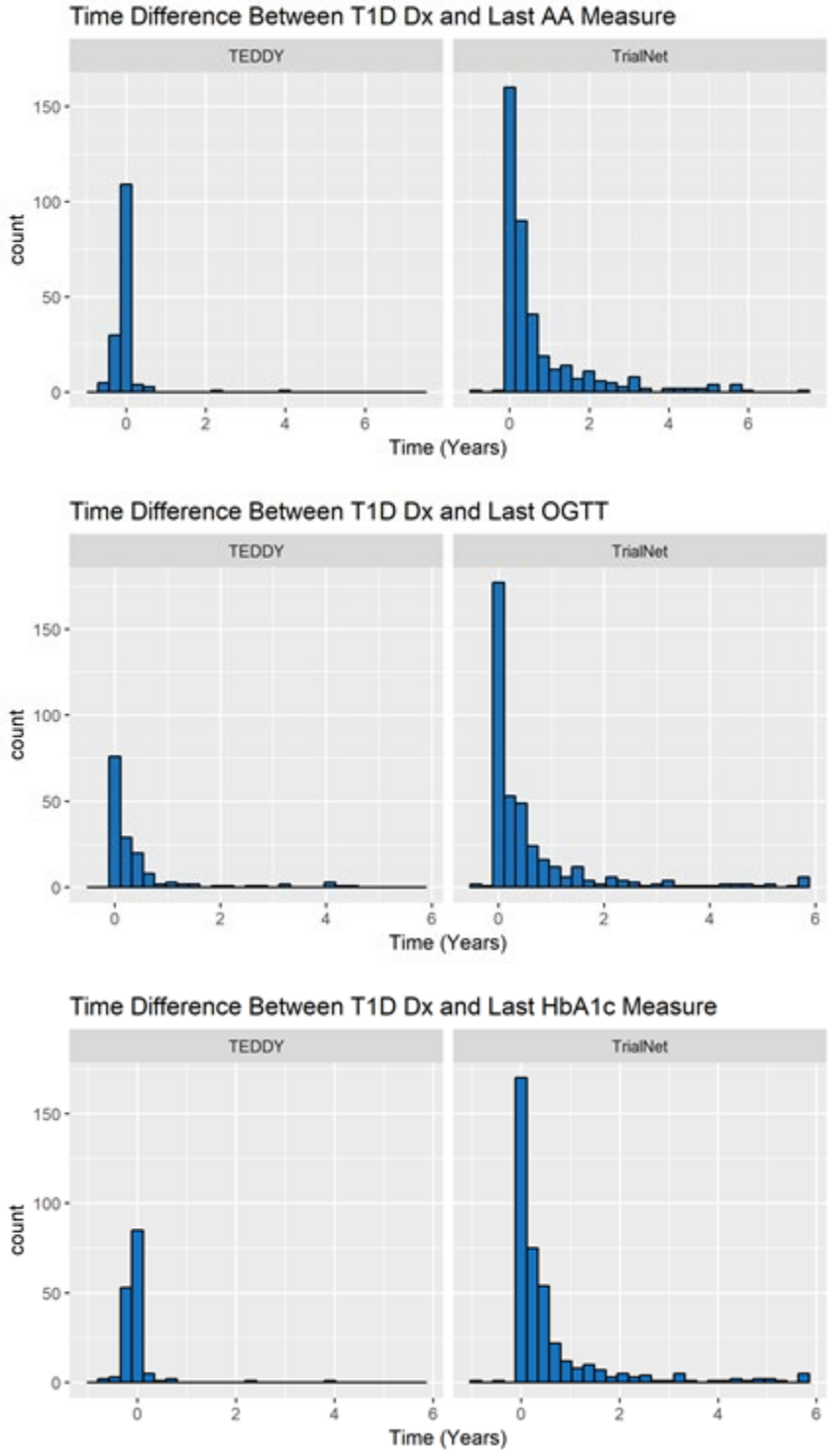
HLA category	TN01	TEDDY	High* or Alternate Risk
DR4*030X/0302*DR3*0501/0201	0	178	High
DR4*030X/0302*DR4*030X/0302	0	80	High
DR4*030X/0302*DR4*030X/020X	2	0	High
DR4*030X/0302*DR8*0401/0402	42	49	High
DR4*030X/0302*DR1*0101/0501	20	6	High
DR4*030X/0302*DR13*0102/0604	53	5	High
DR4*030X/0302*DR4*030X/0304	0	0	High
DR4*030X/0302*DR9*030X/0303	6	0	High
DR3*0501/0201*DR3*0501/0201	81	33	High
DR3*0501/0201*DR9*030X/0303	14	0	High
All others	1070	0	Alternate
Missing	47	2	NA

*High-risk categories used as enrollment criteria for the TEDDY prospective cohort (Hagopian et al. 2011).

4.3.3.3 Assessment of Time-Varying Covariates

The longitudinal information for islet AA positivity and the glycemic markers were assessed for their potential to support a time-to-event model utilizing time-varying covariates. A key assumption of such a model is that the time-varying covariates have measurements at the same time of the event, or within a relatively small difference of time so that it is reasonable to assume the timing of measurements are equal. In the case of the islet AAs, the missing longitudinal information was assessed by measuring the time from the last islet AA measurement to the diagnosis time. The distribution of these times is shown in Figure 2. Between both studies, approximately 50% of the individuals have gaps of timing from their last islet AA measurement to diagnosis over six months. The use of binary time-varying covariates for islet AA positivity would, therefore, require removal of a significant portion of the data. A similar situation arises for OGTT and HbA1c measurements as shown in Figure 2. It is concluded then that the use of time-varying covariates is not feasible.

Figure 2. Distribution of time difference between T1D diagnosis and covariates



4.3.3.4 Assessing Predictors in Individuals with One Islet AA

Whether the population of individuals with one islet AA could be included in the model was assessed by analyzing whether such individuals have features predictive of risk to T1D. For TEDDY there is a complete history of individuals with one islet AA as the subjects are followed from birth. The inclusion criteria for TN01 required individuals to have two or more islet AA before OGTT and HbA1c assessments would be made. Therefore, robust information regarding the glycemic measures for subjects with one islet AA is not available for the modeling purposes outlined in this submission. Table 7 shows diagnosis information for individuals with one islet AA at the earliest record available in the data. The percent of individuals with one islet AA who go onto a diagnosis was 6.5%. The diagnosis event rates for the one islet AA population are low which would make recruitment of these subjects to T1D prevention studies impractical in contrast to those with two or more islet AAs.

Table 7. Diagnosis information for individuals with varying number of islet AAs

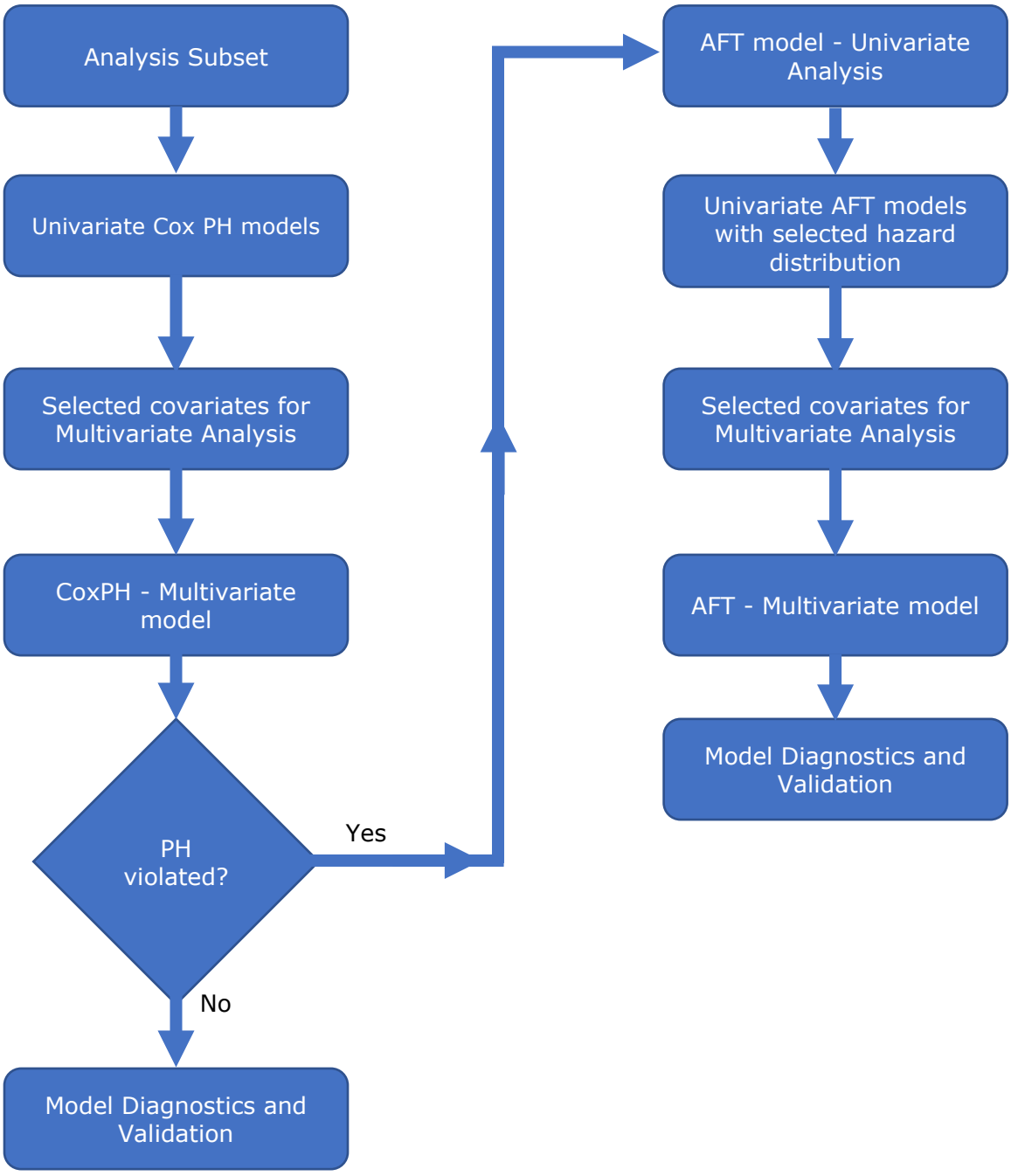
Category	Subjects*	Diagnoses	% Conversion to T1D	Mean Time to Dx (sd)
1 islet AA	9,450	619	6.5	3.09 (2.81)
2 islet AA	2,317	598	25.8	2.16 (2.19)
3 islet AA	1,406	406	28.9	1.92 (2.2)
4 islet AA	635	200	31.5	1.66 (2.07)

*Note the subject numbers captured in t

4.3.4 Modeling Analysis Methodologies

As per the original statistical analysis plan, the first approach was to analyze predictors of T1D diagnosis using a Cox proportional hazard (PH) model, i.e. a semi-parametric approach, as this was the most parsimonious first step. Based on reviewer recommendations, a fully parametric approach was requested. With knowledge of prior quantitative analyses from the literature (Section 4.3.4.1), consideration of the drug development context, and the available data, the full modeling analysis was executed. The flow chart (Figure 3) displays the progression of the modeling analysis, where subsequent steps were executed based on best practices for model building and learnings from previous steps. All analysis was carried out in the R programming language. Specific details of the R packages utilized are highlighted in Section 4.3.4.2. In completion, the model building process followed three main steps: (a) Analysis of Cox PH model using the TN01 and TEDDY datasets and testing the PH assumption; (b) Development of a parametric accelerated failure time model using the TN01 and TEDDY datasets; (c) Evaluation of model performance with k-fold cross-validation and external validation with DAISY as a separate independent dataset (Section 4.3.7).

Figure 3. Modeling development workflow



4.3.4.1 Prior Knowledge

A literature review was conducted to identify previous work using quantitative modeling for predicting T1D diagnosis (Table 8). The highlighted studies focused on quantifying predictors of T1D diagnosis using joint modeling of longitudinal and survival data based on birth cohort data.

Table 8. Historical quantitative modeling for predicting T1D diagnosis

Reference	Model Purpose	Model Description	Data Utilized
(Steck et al. 2015)	Identify and quantify predictors of T1D diagnosis	Survival model with continuous time-varying covariates (islet AA titer values for IA-2, GAD65, and IAA), FDR status, HLA subtype	TEDDY
(Köhler, Beyerlein, et al. 2017)	Quantify the time-varying association between the islet AA titer values and T1D	Bayesian joint modeling of longitudinal and survival data	TEDDY
(Köhler, Umlauf, et al. 2017)	Quantify the time-varying association between the islet AA titer values and T1D	Flexible additive joint modeling of longitudinal and survival data	BABYDIAB and BABYDIET

4.3.4.2 Software

Model building, visualization, model assumptions, diagnostics and external validation was conducted in R (version 4.0.0; Vienna, Austria, R Core Team, 2018) using the packages "survival" (Therneau 2020), "flexsurv" (Jackson 2016), "survminer" (Kassambara and Kosinski 2018), "dplyr" (Wickham et al. 2020), "survAUC" (Potapov, Adler, and Schmid 2015), "rms" (Harrell 2019), survParamSim (Yoshida and Claret 2020) and "riskRegression" (Ozenne et al. 2017).

4.3.5 Cox Proportional Hazard Model

The semiparametric Cox PH model relates the T1D diagnosis events with the covariates,

$$h_i(t) = h_0(t)\exp\left(\sum_{j\in I}\beta_j X_{ij}\right) \quad \text{(E1)}$$

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 $h_0(t)$ is the baseline hazard. The use of a Cox PH model implies that the underlying baseline hazard function is not specified to have a parametric distribution and that the PH assumption holds, i.e. the ratio of hazards between different individuals remains constant over time.

4.3.5.1 Univariate Analysis

A univariate analysis was performed by estimating a Cox PH model for of the covariates in Table 3. The 'coxph' function in the 'survival' R package was used for Cox PH analysis (Therneau 2020). Covariates with no significant univariate association (P-value ≥ 0.1) with T1D diagnosis were not considered for the full model development. The p-value was computed using the Wald test, which evaluates whether the covariate coefficient is statistically different from zero. A multiplicity adjusted alpha value (Bonferroni correction) was used for univariate analysis.

4.3.5.2 Analysis of Correlation and Association between Covariates

The covariates remaining after the univariate analysis were analyzed for multicollinearity and associations prior to performing multivariate analysis. Pearson's correlation was used to test the correlation between continuous covariates, with a correlation value above 0.3 chosen as significant. The Wilcoxon test was used to test the association between continuous and categorical covariates, and the Chi-square test of independence was used to test the association between categorical covariates. In both cases, a p-value < 0.001 (multiplicity adjusted) was chosen as the threshold for significance.

4.3.5.3 Multivariate Analysis

The multivariate analysis was performed by testing all possible combinations of remaining covariates, as the number of covariates for multivariate analysis were reasonable. The comparison between possible models was conducted using Akaike's Information Criteria (AIC). A reduction in AIC value greater than or equal to 10 suggests a strong evidence in favor of the model with lower AIC (Burnham and Anderson 2016).

4.3.5.4 Model Diagnostics

To assess if the PH assumption was satisfied, Schoenfeld residuals were utilized. The expected value of these residuals can be used to quantify potential time-dependency on survival times. The Pearson product-moment correlation between the scaled Schoenfeld residuals and log(time) for each covariate was computed using the 'cox.zph' function in R. Values below a significance threshold indicated a violation of the PH assumption. Additional model diagnostics were not performed for the Cox PH model due to a violation of the PH assumption observed with the above-mentioned test.

4.3.6 Parametric Accelerated Failure Time Model

The AFT model was chosen as the modeling methodology after assessing the Cox PH model because it does not require satisfaction of the PH assumption. It assumes that the effect of a covariate is to adjust (accelerate or decelerate) the time course of the event of interest and is given by (Jackson 2016),

$$S_i(t) = S_0\left(t \exp\left(-\sum_{j \in I} \beta_j X_{ij}\right)\right) \quad \text{(E2)}$$

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 9 provides the survival functions for a list of different forms of parametric distribution. The β-parameter value specifies the effect each covariate has on the survival time, where negative β values indicate that the survival time increases with positive-valued covariates, and positive β values indicate that the survival time decreases with positive-valued covariates. For 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 9. For R-package flexsurvreg, the output "scale" is the scale (λ) and the output "shape" is the shape (α) parameter as shown in Table 9.

4.3.6.1 Selection of Parametric Distribution

Multiple parametric distributions were tested for their ability to approximate the underlying hazard function including exponential, Weibull, gamma, generalized gamma, generalized F, log logistic, log normal and Gompertz. Resulting Akaike information criterion (AIC) values and graphical methods for survival and hazard function fits were compared to select an appropriate parametric form. The 'flexsurvreg' function in the 'flexsurv' R package was used for the selection of parametric distribution analysis.

Table 9. 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.

4.3.6.2 Univariate Analysis

A univariate analysis was performed by estimating an AFT model using the parametric distribution selected from Section 4.3.6.1, for each of the covariates in Table 3. The 'flexsurvreg' function in the 'flexsurv' R package was used to perform parametric AFT model analysis. Individual covariates with no significant association (P-value ≥ 0.05) with T1D diagnosis were not considered for the full model development. The p-value was computed using the Wald test, as described in Section 4.3.5.1. A multiplicity adjusted alpha value (Bonferroni correction) was used for univariate analysis. The remaining covariates were analyzed for multicollinearity and associations prior to performing multivariate analysis.

4.3.6.3 Analysis of Correlation and Association between Covariates

The analysis defined in Section 4.3.5.3 was repeated for the covariates remaining after the AFT univariate analysis.

4.3.6.4 Multivariate Analysis

The multivariate analysis was performed by testing all possible combinations of covariates, as the number of covariates for multivariate analysis were reasonable. The comparison between possible models was made using AIC criteria. A reduction in AIC value greater than or equal to 10 suggests strong evidence in favor of the model with lower AIC (Burnham and Anderson 2016). Additionally, per EMA SAWP feedback, three alternative models were added to the list of multivariate models. These models included baseline age and sex as covariates in different combinations, in addition to the previous covariates, included from the multivariate analysis.

4.3.6.5 Model Diagnostics

Quantile-Quantile (Q-Q) plots were used to assess the validity of the AFT model assumption for two groups of survival data. In this case, such groups correspond to the presence or absence of an AA combination. Under the AFT model assumption, the presence of one islet AA combination has a multiplicative effect on survival time. Conceptually, a Q-Q plot examines various percentiles for which the survival times are computed for the two groups. A plot of the survival times for the chosen percentiles should give a straight line if the AFT model is appropriate, where the straight line is an estimate of the acceleration factor. Such plots were generated for each AA combination in the AFT model. To analyze continuous covariates, binary groups were formed using thresholds to allow for the generation of Q-Q plots.

4.3.7 Model Performance and Validation

4.3.7.1 Model Performance

To assess the model's predictive performance on the analysis set, time-dependent receiver operating characteristic (ROC) curves were generated (Heagerty and Zheng 2005). Conceptually, the methodology of this metric is that model predictions on all at-risk individuals up to a time t are derived, and true/false positive rates based on model predictions versus the observed data are computed. This is repeated across multiple timepoints to generate ROC curves. The area under the ROC curves (AUC) are computed, which are interpreted as the concordance between the model prediction and data. This methodology is an appropriate model performance metric as an individual's risk for developing T1D changes over time. Further, it provides metrics as to the model's predictive power for time frames over which a trial of reasonable duration would be conducted. To further explore performance, predictions on each fold stratified by individual covariates were performed. Goodness-of-fit plots (VPC-style) were created for visual assessments of models fits (Appendix G). The 'survParamSim' R package was used to generate the VPC-style plots. 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.

4.3.7.2 K-fold Cross-Validation

Model validation was performed using the k-fold cross-validation technique (Breiman and Spector 1992). Data was 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, while the remaining were used as the training set for all combinations. Goodness-of-fit plots were created by overlaying the model estimated survival on Kaplan-Meier curves for all five folds. The concordance index was computed for each of the five folds estimated by time increments of one year up to six years.

4.3.7.3 Cross-Validation on Pediatric Population

An internal validation was performed by analyzing predictive performance on pediatric subpopulations in the data. A randomly selected portion (50%) of individuals aged less than an age threshold was extracted and used as a test data set. The remaining data constituted the training data used to fit the model. Goodness-of-fit plots were created by overlaying model estimated survival on Kaplan-Meier curves. The concordance index was computed for time increments of one year up to six years.

4.3.7.4 External Validation

External validation was performed using the DAISY dataset described in Section 4.3.1. The definition of the derived baseline defined in Section 4.3.2.1 was applied to the data to arrive at a validation set. The AFT model was used to predict survival within this subset. Goodness-of-fit (VPC style) plots were created by overlaying model estimated survival on Kaplan-Meier curves. The concordance index was computed for time increments of one year up to six years.

4.4 Results

The execution of the modeling methodologies outlined in Section 4.3 are reported here. The Cox PH model focused on checking the validity of the PH assumption. As this assumption was not met, a parametric AFT model was chosen. Model diagnostic, performance, and validation exercises were performed to assess the model's ability to quantify the time-varying effect of islet AAs and glycemic markers on risk to T1D diagnosis. For ease of viewing, the notation for covariates is shown using the underlying subscript, e.g. $X_{\text{GAD65\_IAA}}$ is equivalent to GAD65_IAA.

4.4.1 Cox Proportional Hazard Model

4.4.1.1 Univariate Analysis

A univariate analysis was performed using the covariates listed in Table 3. Estimates of each individual regression beta coefficient, effect size (hazard ratio), and statistical significance was determined with respect to the overall survival (Table 10). The covariates SEX, bAGE_s, 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 had statistically significant beta coefficients. The covariates for 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, GAD65_IAA_IA-2) did not show a significant effect on overall survival and were dropped from subsequent analysis based on a p-value < 0.1 (multiplicity adjusted). The negative beta coefficient for some islet AA combinations, such as GAD65_IAA indicates lower hazard relative to the baseline hazard function, which accounts for risk of all other combinations of AA. All the covariates that showed statistical significance were considered for multivariate analysis.

Table 10. Univariate analysis for each covariate using Cox PH Model

Covariate	beta	HR (95% CI)	Wald statistic	p-value	Significant
TEDDY_Trial	-0.001	1 (0.82-1.2)	0	0.99	No
SEX	-0.26	0.77 (0.65-0.91)	9	0.0026	Yes
bAGE_s	-0.26	0.77 (0.69-0.85)	24	8.40E-07	Yes
BMI	-0.027	0.97 (0.95-1)	4	0.044	No
HR_HLA	0.093	1.1 (0.92-1.3)	1.1	0.3	No
FDR	-0.011	0.99 (0.8-1.2)	0.01	0.92	No
GAD65_IAA	-0.71	0.49(0.37-0.65)	24	1.20E-06	Yes
GAD65_ZnT8	-0.8	0.45 (0.32-0.62)	23	1.40E-06	Yes
GAD65_IA-2	0.07	1.1 (0.8-1.4)	0.22	0.64	No
IA-2_IAA	0.4	1.5 (0.8-2.8)	1.6	0.21	No
IA-2_ZnT8	0.75	2.1 (1.5-2.9)	19	1.30E-05	Yes
IAA_ZnT8	-0.073	0.93 (0.5-1.7)	0.05	0.82	No
GAD65_IA-2_IAA	0.2	1.2 (0.84-1.8)	1.1	0.29	No
GAD65_IAA_ZnT8	-0.27	0.76 (0.55-1.1)	2.5	0.11	No
GAD65_IA-2_ZnT8	0.14	1.2 (0.92-1.4)	1.6	0.21	No
IA-2_IAA_ZnT8	0.72	2.1 (1.5-2.9)	18	2.00E-05	Yes
GAD65_IA-2_IAA_ZnT8	0.44	1.6 (1.3-1.9)	18	2.00E-05	Yes
Log_GLU120_s	0.78	2.2 (2-2.4)	240	9.40E-55	Yes
Log_GLU0_s	0.19	1.2 (1.1-1.3)	15	8.30E-05	Yes
HbA1c_s	0.56	1.7 (1.6-1.9)	130	1.10E-30	Yes

4.4.1.2 Analysis of Correlation and Association between Covariates

The correlation between the continuous covariates (Figure 4) did not reveal any covariate pairs with high correlation, defined as correlations above 0.3. The Wilcoxon test (Table 11) and the chi-square test of independence (Table 12) showed that the baseline Age (bAGE_s) and SEX were highly associated with AA combinations. Hence, bAGE_s and SEX were dropped from subsequent analysis. Association between islet AA combinations was not considered relevant as their presence is mutually exclusive i.e. only one islet AA combination is possible for a given subject. Hence, based on the correlation and association analysis between

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 multivariate analysis.

Figure 4. Pearson's correlation between continuous covariates

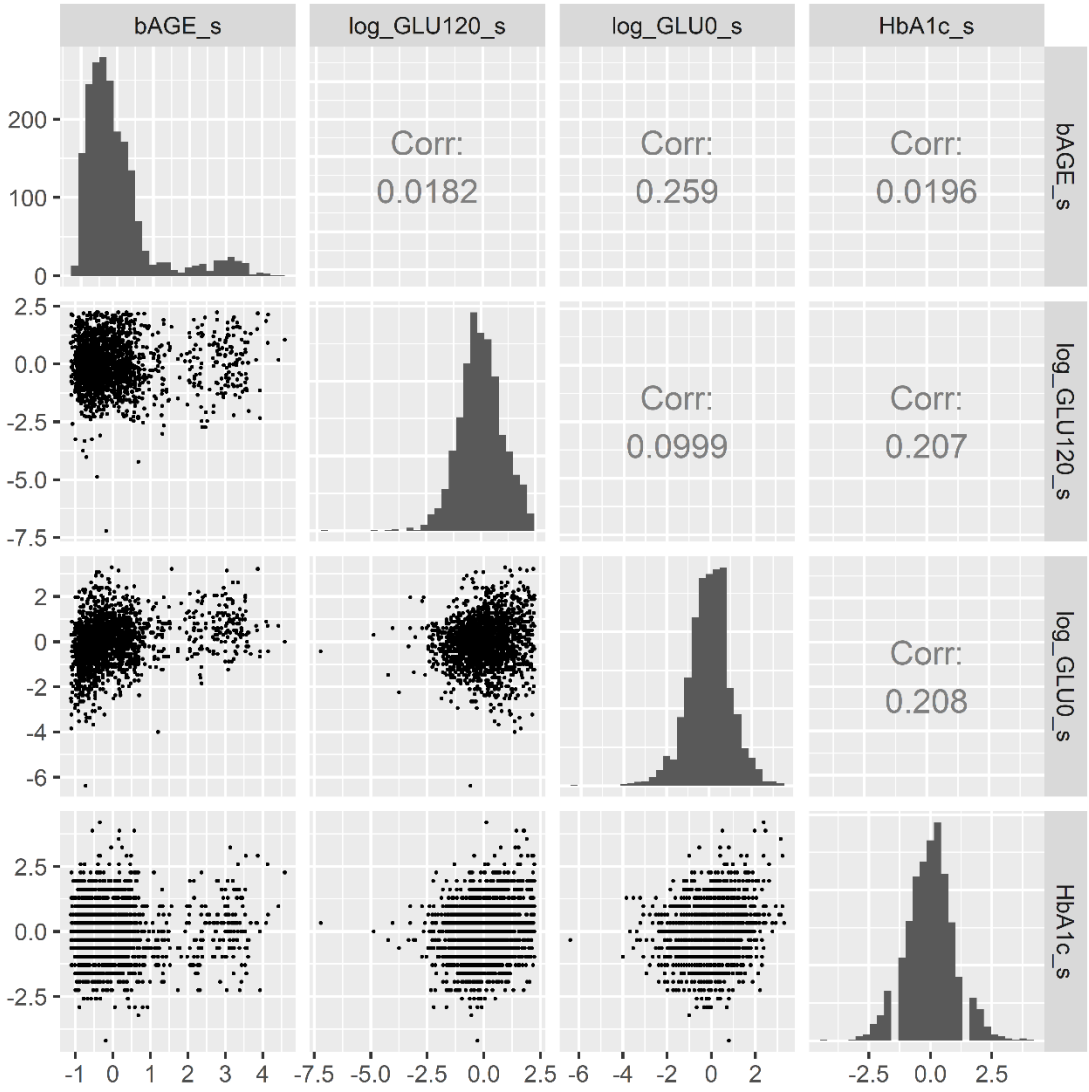


Table 11. Wilcoxon test between continuous and categorical covariates

Covariate	SEX	GAD65_ IAA	GAD65_ ZnT8	IA-2_ ZnT8	IA-2_ IAA_ZnT8	GAD65_IA-2 _IAA_ZnT8
bAGE_s	1.28E-02	3.31E-07	1.05E-16	3.51E-01	2.81E-10	1.14E-07
Log_GLU120_s	9.26E-02	7.38E-03	2.17E-03	3.76E-03	1.31E-03	5.45E-02
Log_GLU0_s	2.60E-04	6.85E-01	2.67E-01	2.29E-01	5.58E-01	4.10E-01
HbA1c_s	1.56E-01	4.37E-01	1.05E-01	2.30E-01	1.36E-01	7.22E-02

Table 12. Chi-square test of independence between categorical covariates

	GAD65_ IAA	GAD65_ ZnT8	IA-2_ ZnT8	IA-2_ IAA_ZnT8	GAD65_IA-2_ IAA_ZnT8
SEX	7.55E-01	4.07E-02	6.57E-05	4.13E-03	7.96E-01

4.4.1.3 Multivariate Analysis

A total of eight possible multivariate models were obtained based on covariates selected from univariate analysis (Section 4.4.1.1) and analysis of correlation and association Section 4.4.1.2. Islet AA combinations remaining after the univariate analysis were included to generate the base model as this multi-level covariate was of primary interest. AIC values were computed for these eight model fits (Table 13). Models 6 and 8 produced the least AIC values. Additionally, AIC values for these two models were significantly lower (> 10) compared to all other models. Among these two models, model 6 was selected as the selected Cox PH model as it produced a lower AIC with a lower number of covariates. In summary, a Cox PH model and covariates GAD65_IAA, GAD65_ZnT8, IA-2_ZnT8, IA-2_IAA_ZnT8, GAD65_IA-2_IAA_ZnT8, Log_GLU120_s and HbA1c_s was selected as the selected multivariate model. Table 14 provides the parameter estimates for the selected model. Appendix H Table 1-7 provides parameter estimates for other models listed in Table 13.

Table 13. Value of AIC for models fitted with Cox PH

Model	Covariates	AIC
1	GAD65_IAA + GAD65_ZnT8 + IA-2_ZnT8 + IA-2_IAA_ZnT8 + GAD65_IA-2_IAA_ZnT8 (Base model)	7038.28
2	Base model + Log_GLU0_s	7024.95
3	Base model + HbA1c_s	6918.96
4	Base model + Log_GLU120_s	6808.17
5	Base model + Log_GLU120_s + Log_GLU0_s	6801.93
6	Base model + Log_GLU120_s + HbA1c_s	6730.80
7	Base model + Log_GLU0_s + HbA1c_s	6918.23
8	Base model + Log_GLU0_s + Log_GLU120_s + HbA1c_s	6732.35

Table 14. Selected Cox PH model parameter estimates (model 6)

Covariate	beta	Std Error (beta)	HR	Wald Statistic	p-value
GAD65_IAA	-0.58258	0.15354	0.55845	-3.794	0.000148
GAD65_ZnT8	-0.7244	0.17331	0.48461	-4.18	2.92E-05
IA-2_ZnT8	0.41654	0.17858	1.5167	2.333	0.019673
IA-2_IAA_ZnT8	0.4713	0.17536	1.60207	2.688	0.007196
GAD65_IA-2_IAA_ZnT8	0.19173	0.11255	1.21134	1.704	0.088463
Log_GLU120_s	0.69255	0.05145	1.99881	13.46	< 2e-16
HbA1c_s	0.41789	0.04699	1.51875	8.893	< 2e-16

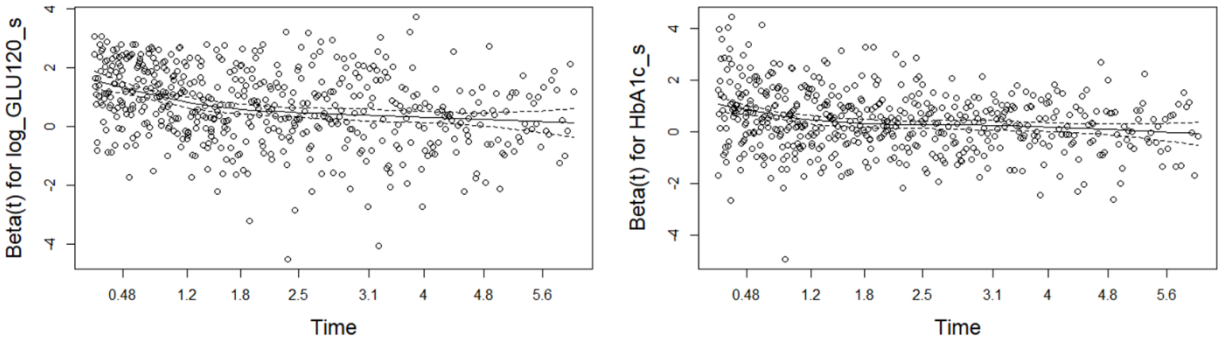
4.4.1.4 Model Diagnostics

The model diagnostic to test the PH assumption was performed using the 'cox.zph function' on the selected Cox PH model described in Section 4.4.1.3. The results showed that HbA1c_s and Log_GLU120_s violated the proportional hazards assumption (Table 15). The global p-value was also less than 0.05. To further verify this result, the scaled Schoenfeld residual plot showed a systematic departure from the horizontal line for HbA1c_s and GLU120_s indicating a clear dependency on time, i.e. a violation of proportional hazards assumption (Figure 5). The Schoenfeld residuals plots for islet AA combination in the selected model are available in Appendix H Figure 27-31. For subsequent analysis, an AFT model was analyzed as it does not require the PH assumption to hold.

Table 15. Testing proportional hazards assumption using Schoenfeld residuals to test for independence between residuals and survival time

Covariate	p-value
GAD65_IAA	0.74
GAD65_ZnT8	0.31
IA-2_ZnT8	0.81
IA-2_IAA_ZnT8	0.47
GAD65_IA-2_IAA_ZnT8	0.98
Log_GLU120_s	<2.0 E-16
HbA1c_s	1.6 E-10
GLOBAL	<2.06E-16

Figure 5. Graphical diagnostics with scale Schoenfeld residuals (Beta(t)) against survival time



4.4.2 Parametric Accelerated Failure Time Model

4.4.2.1 Selection of Parametric Distribution

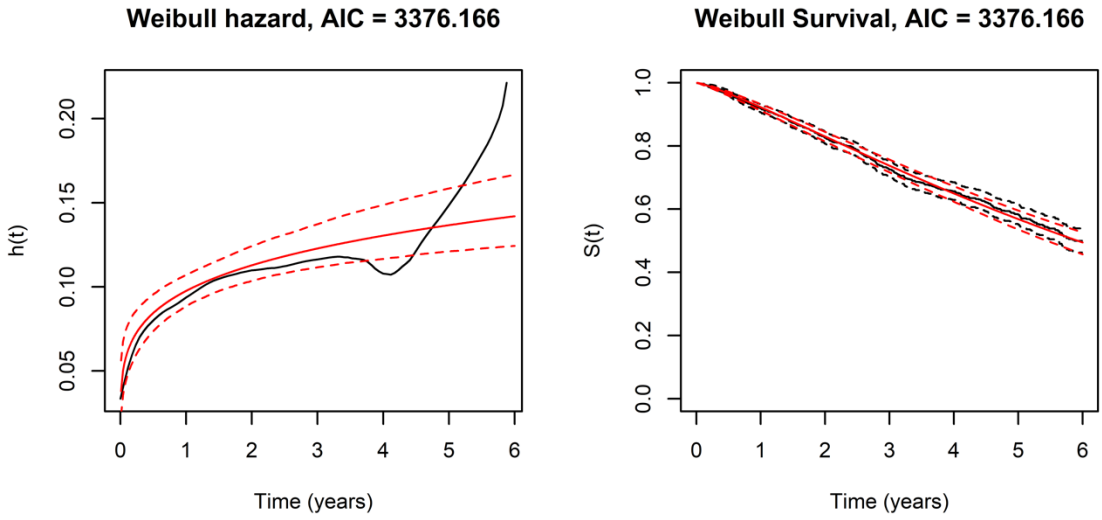
Several parametric distributions were tested as the basis for the probability density in the AFT model. The Weibull distribution was found to be the most appropriate distribution to parameterize the form of hazard function based on AIC and graphical inspection (Table 16, Figure 6). Large deviations of the estimated hazard function represented by the data (black line in Figure 6) are seen in comparison to the Weibull profile (red line in Figure 6) for times beyond four years. This is due to the sparseness of data for longer diagnosis time but does not have a large impact on survival prediction (see survival estimate in Figure 6). The AIC

values were computed using the base AFT model, i.e. without covariates (null model). Weibull, gamma, generalized gamma, generalized F, log logistic distributions were seen to have similar AIC values indicating that any choice from among them is immaterial for improved model fitting. The Weibull distribution was selected among these distributions based on graphical inspection for survival and hazard functions (Figure 6). Additionally, the Weibull distribution is well-characterized and has several supporting software packages. The hazard function and cumulative hazard plots for other distributions are available in Appendix H Figure 32-38.

Table 16. AIC value for hazard distribution

Hazard Distribution	AIC
Exponential	3400.20
Weibull	3376.17
Gamma	3374.81
generalized gamma	3375.76
generalized F	3377.76
log logistic	3375.35
Gompertz	3387.68
log normal	3392.20

Figure 6. Survival and hazard plots for Weibull Distribution. (Black lines represent the estimated hazard and survival functions from the data, and red solid and dotted lines indicate the mean and 95% confidence intervals of the Weibull model fit.)



4.4.2.2 Univariate Analysis

A univariate AFT model analysis using the Weibull distribution was performed using the covariates listed in Table 3. The beta coefficients, 95% confidence interval for the coefficient, and statistical significance for each of the covariate with respect to the overall survival is shown in (Table 17). The covariates bAGE_s, 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 had statistically significant beta coefficients. The covariates SEX, 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, GAD65_IAA_IA-2) did not show significant effect on

overall survival and were dropped from subsequent analysis. All the covariates that showed statistical significance ($p < 0.05$, multiplicity adjusted) were considered for subsequent analysis.

Table 17. Univariate analysis for each covariate using AFT model with Weibull distribution

Covariate	beta	95% lower CI	95% upper CI	p-value	Significant
TEDDY_Trial	0.0109	-0.151	0.173	0.895	No
SEX	0.218	0.0755	0.361	0.00273	No
bAGE_s	0.217	0.129	0.306	1.56E-06	Yes
HR_HLA	-0.0684	-0.213	0.0765	0.355	No
FDR	-0.00096	-0.175	0.173	0.991	No
BMI	0.0212	0.000217	0.0421	0.0477	No
GAD65_IAA	0.587	0.348	0.826	1.50E-06	Yes
GAD65_ZnT8	0.663	0.392	0.935	1.66E-06	Yes
GAD65_IA-2	-0.0571	-0.298	0.184	0.643	No
IA-2_IAA	-0.329	-0.846	0.189	0.214	No
IA-2_ZnT8	-0.614	-0.892	-0.337	1.40E-05	Yes
IAA_ZnT8	0.0653	-0.452	0.583	0.805	No
GAD65_IA-2_IAA	-0.163	-0.473	0.147	0.303	No
GAD65_IAA_ZnT8	0.221	-0.056	0.498	0.118	No
GAD65_IA-2_ZnT8	-0.117	-0.299	0.0656	0.209	No
IA-2_IAA_ZnT8	-0.592	-0.868	-0.316	2.57E-05	Yes
GAD65_IA-2_IAA_ZnT8	-0.368	-0.536	-0.199	1.91E-05	Yes
Log_GLU120_s	-0.607	-0.687	-0.526	2.07E-49	Yes
Log_GLU0_s	-0.156	-0.232	-0.0789	7.01E-05	Yes
HbA1c_s	-0.449	-0.529	-0.369	5.08E-28	Yes

4.4.2.3 Analysis of Correlation and Association

The correlation and association analysis are irrespective of the choice of the survival model; hence, the analysis described in Section 4.4.1.2 was used to assess covariate selection for AFT multivariate analysis. As a result, based on AFT univariate analysis (Section 4.4.2.2) and Analysis of Correlation and Association (Section 4.4.1.2) 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.

4.4.2.4 Multivariate Analysis

A total of eight possible multivariate models were obtained based on covariates selected from univariate analysis (Section 4.4.2.2) and analysis of correlation and association (Section 4.4.2.3). All significant AA combinations from univariate analysis were included to generate the base model as these covariates were of primary interest. Models 6 and 8 produced the least AIC values (Table 18). Additionally, AIC values for these two models were significantly lower (> 10) compared to all other models. Among these two models, model 6 produced a lower AIC with a lower number of covariates. In summary, an AFT model with Weibull distribution and covariates GAD65_IAA, GAD65_ZnT8, IA-2_ZnT8, IA-2_IAA_ZnT8,

GAD65_IA-2_IAA_ZnT8, Log_GLU120_s and HbA1c_s was chosen (model 6) among the 8 models. Table 19 provides the parameter values for model 6 with shape and scale parameters for the Weibull distribution, estimated beta values, and Wald test p-values for each covariate. Appendix H Table 8-14 provides parameter estimates for other models listed in Table 18. As per EMA SAWP feedback, alternative models were developed by adding bAGE_s and SEX in different combinations to model 6 (Table 20). For comparison, the originally proposed model 6 will be referred to as the original model (orig_mod). The original model AIC value was compared with the alternative models (Table 20). 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. Model performance and validation were executed for the selected model (alt_mod3) as discussed in the subsequent sections. Table 21 shows the parameter estimates for the selected model (alt_mod3).

Table 18. 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)	3292.476
2	Base model + Log_GLU0_s	3278.769
3	Base model + HbA1c_s	3173.157
4	Base model + Log_GLU120_s	3059.067
5	Base model + Log_GLU120_s + Log_GLU0_s	3052.591
6	Base model + Log_GLU120_s + HbA1c_s	2981.886
7	Base model + Log_GLU0_s + HbA1c_s	3172.244
8	Base model + Log_GLU0_s + Log_GLU120_s + HbA1c_s	2983.369

Table 19. Model 6 (orig_mod) parameter estimates

Covariates	Beta	95% lower CI	95% upper CI	p-value
Shape	1.350	1.260	1.440	NA
Scale	7.710	6.901	8.634	NA
GAD65_IAA	0.434	0.210	0.659	1.50E-04
GAD65_ZnT8	0.539	0.286	0.792	2.95E-05
IA-2_ZnT8	-0.303	-0.562	-0.043	2.21E-02
IA-2_IAA_ZnT8	-0.342	-0.597	-0.086	8.69E-03
GAD65_IA-2_IAA_ZnT8	-0.143	-0.306	0.021	8.78E-02
Log_GLU120_s	-0.518	-0.594	-0.441	5.64E-40
HbA1c_s	-0.309	-0.379	-0.239	3.42E-18

Table 20. 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	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 21. 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	-0.087	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

4.4.2.5 Model Diagnostics

The diagnostic analysis indicated that the AFT models adequately described the effect of AA combination status (presence or absence), HbA1c_binary, AGE_binary, SEX, and GLU120_binary. The Q-Q plots for each covariate produced points that approximate a straight line through the origin suggesting the validity of the AFT model (Figure 7).

Figure 7. QQ plots for categorical covariates in the model

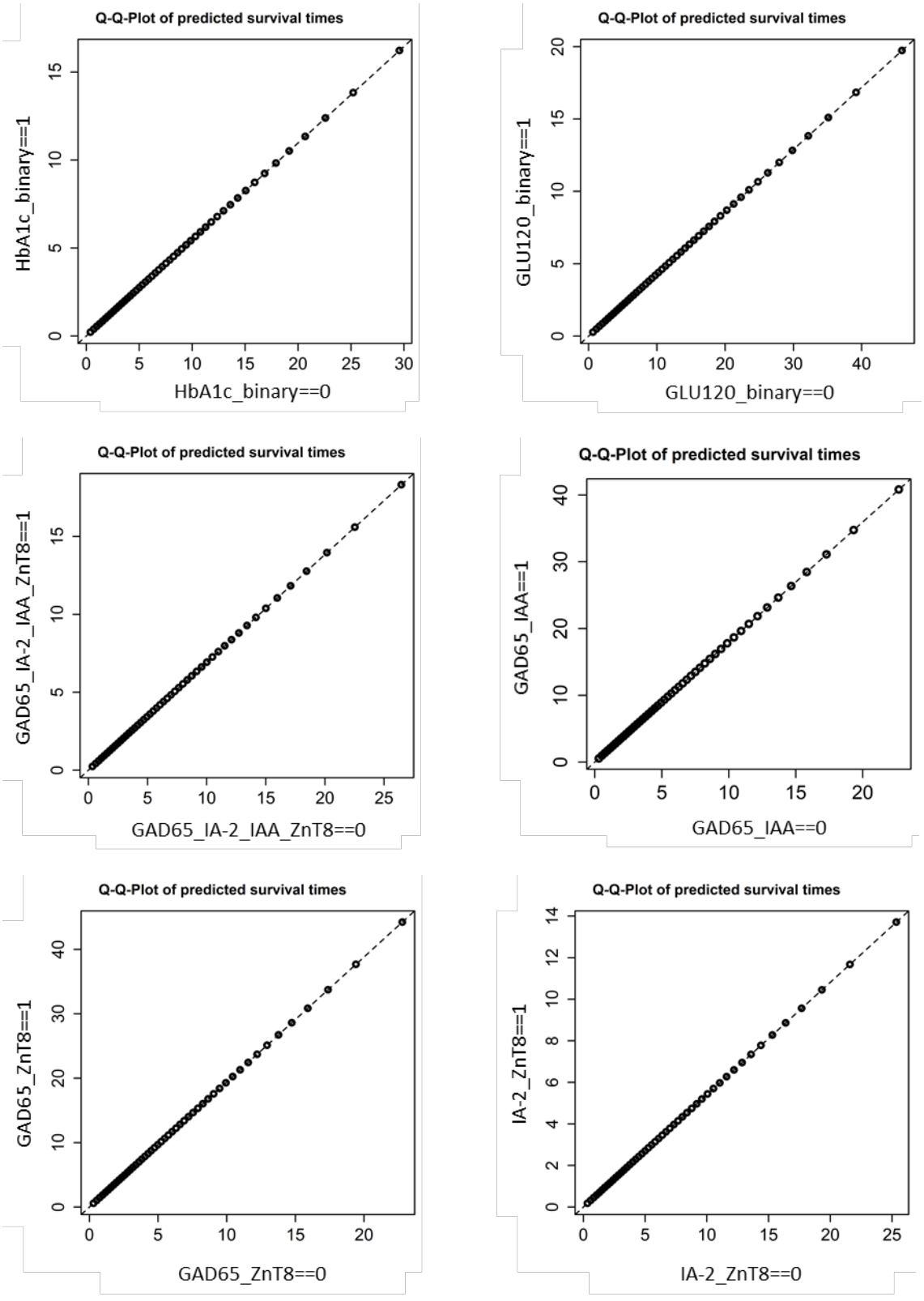
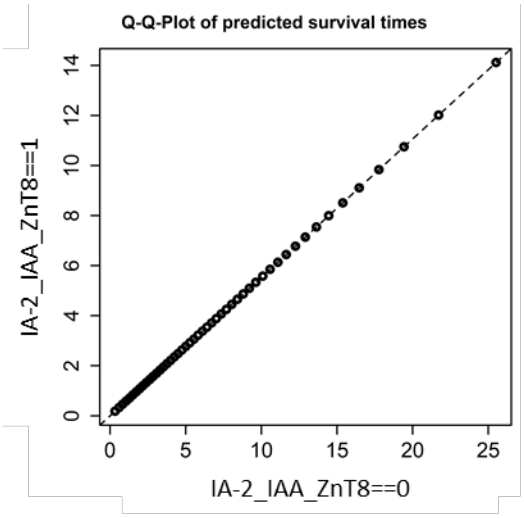
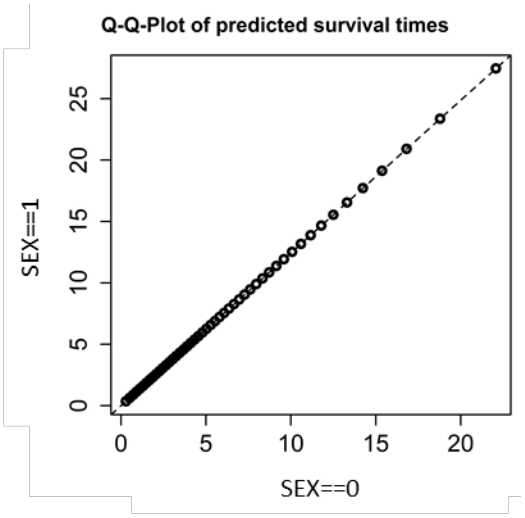
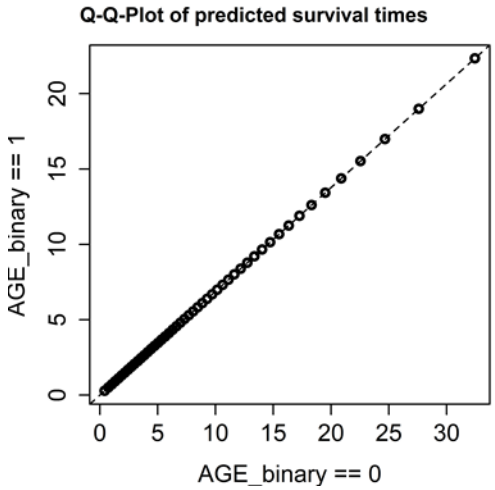


Figure 7 (continued).





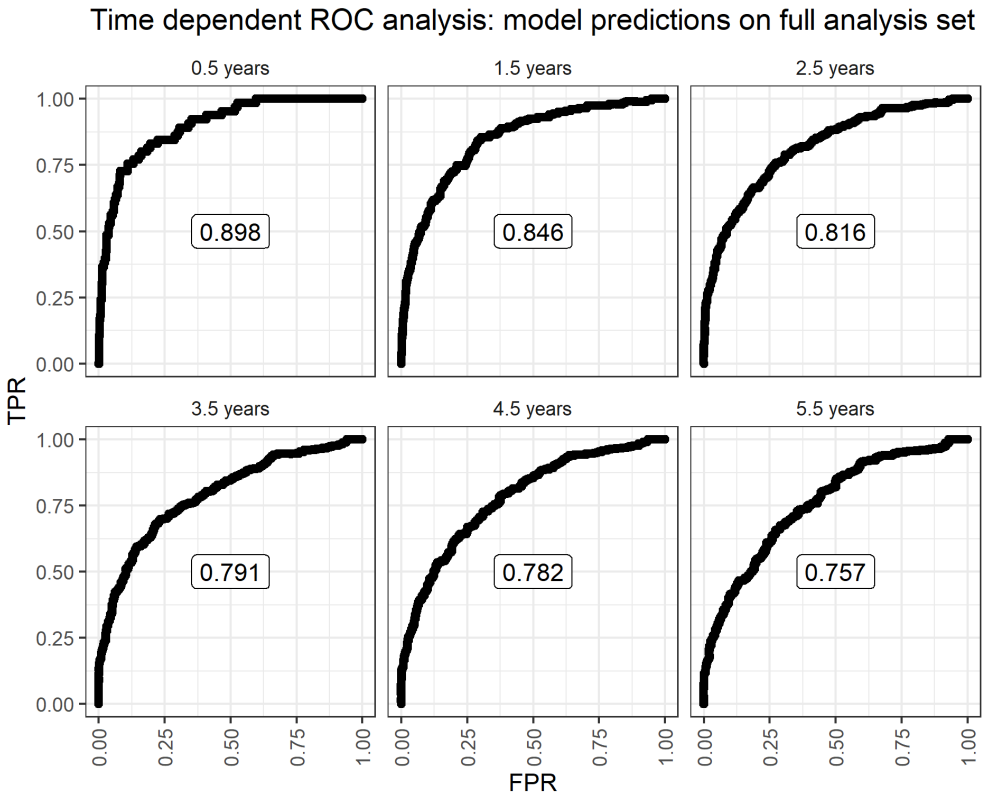


4.4.3 Model Performance and Validation

4.4.3.1 Model Performance

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 8). 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.

Figure 8. Evaluation of model performance using time dependent receiver operation characteristic (ROC) analysis



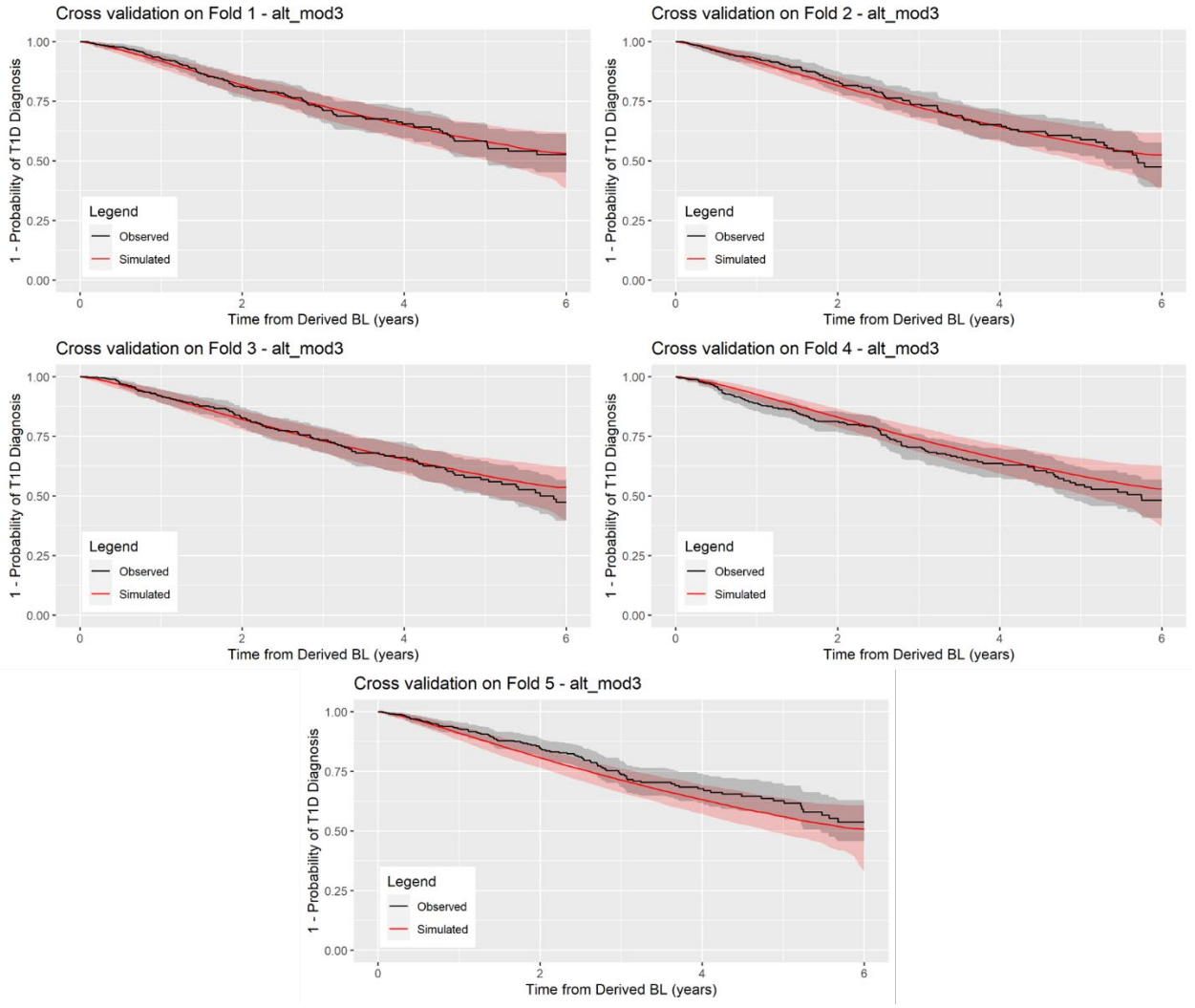
4.4.3.2 K-fold Cross Validation

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 22). 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 9). The black curve represents the Kaplan–Meier estimate, and the red curve represents model prediction. These results provide evidence for good predictive power for time frames over which clinical trials of reasonable duration would be conducted. To assess performance across the covariates, plots were created to show model predictions stratified by each of the islet AA combinations and continuous covariates using binary groups (Appendix H Figure 39-73).

Table 22. 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 9. VPC-style plots for k-fold cross validation (red shaded region shows the 95% prediction interval and the black shaded region shows the 95% CI for the observed data)



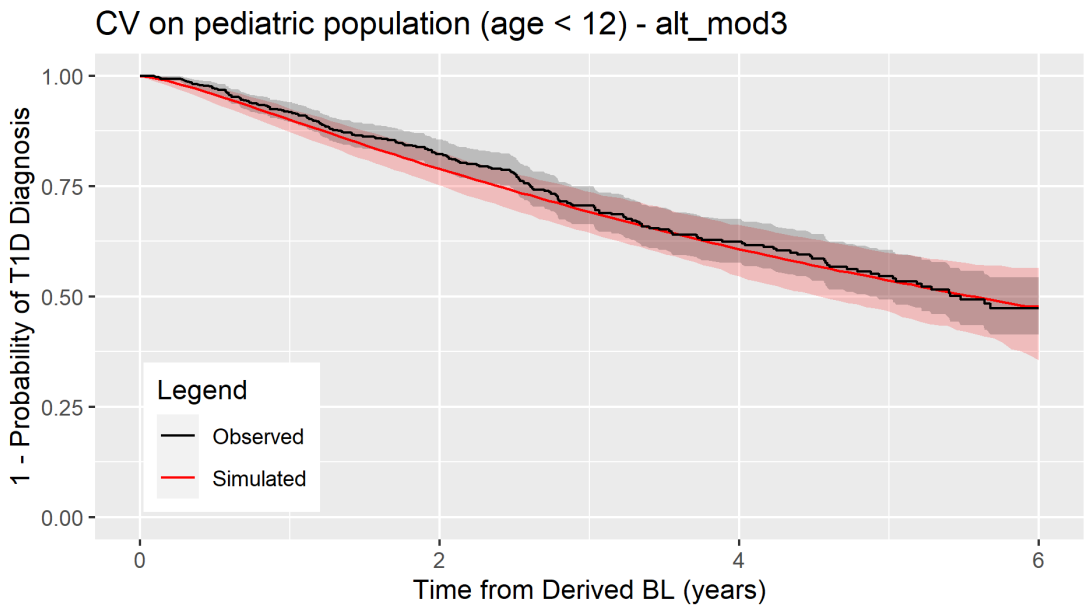
4.4.3.3 Cross-Validation on a Pediatric Population (age < 12)

A pediatric population (age < 12) was derived in the analysis dataset comprised of 1330 subjects, with 345 from TEDDY and 985 from TN01. Half of this population, i.e. 665, were randomly selected as a test set for this cross-validation analysis. 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 23). The visual predictive check (VPC) performed on the survival plot for cross-validation on the pediatric population (age < 12) showed reasonable graphical fit (Figure 10). The black curve represents the Kaplan–Meier estimate for the observed data, and the red curve represents the model prediction. The median of the prediction interval was within the 95% CI band of the estimated Kaplan-Meier curve for the observed data.

Table 23. C-index values over six years with cross-validation on a pediatric population (age < 12)

	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 10. VPC-style plot for internal cross validation (CV) using pediatric population (red shaded region shows the 95% prediction interval and the black shaded region shows the 95% CI for the observed data)



4.4.3.4 External Validation

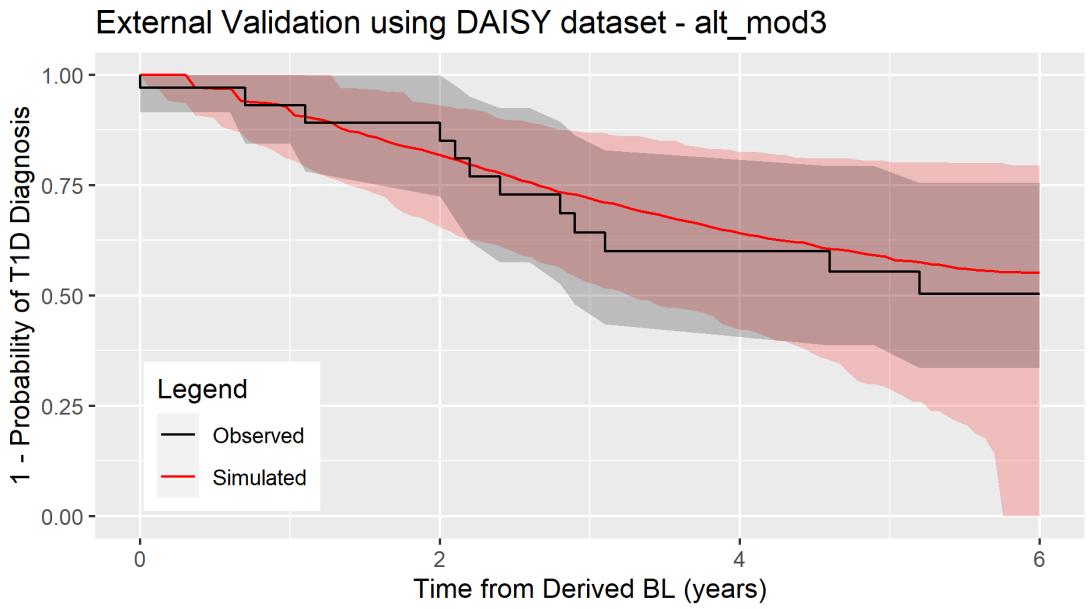
For external validation with DAISY dataset, the selected model (alt_mod3) achieved a c-index 0.91 and 0.82 in years one and two, respectively, even with a limited number of subjects, 34 in the external dataset (Table 24). However, the c-index values beyond three years were relatively lower than up to 2 years, likely attributable to the sparsity of T1D diagnoses during the later years in the DAISY analysis set (Table 24). The VPC performed on the survival plot showed good graphical fit given the limited number of events (Figure 11). These results

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

Table 24. C-index values over six years with DAISY external validation dataset

C-index	Up to year 1	Up to year 2	Up to year 3	Up to year 4	Up to year 5	Up to year 6
alt_mod3	0.91	0.82	0.67	0.68	0.67	0.66

Figure 11. VPC-style plot for external validation using the DAISY analysis dataset (red shaded region shows the 95% prediction interval and the black shaded region shows the 95% CI for the observed data)



4.5 Intended Application of Proposed Tool

The islet AAs are intended to be leveraged as enrichment biomarkers as a means of patient selection in clinical trials investigating therapies that are intended to prevent or delay the clinical diagnosis of T1D. These biomarkers, along with additional patient features, including baseline HbA1c and the 120-minute timepoint from an OGTT, can be used as predictors to identify subpopulations at highest risk of a diagnosis of T1D during T1D prevention clinical trials.

4.5.1 Methodology of Tool

The proposed model is intended to be the foundation of a fully functioning end-user tool that will allow sponsors to optimize enrichment criteria for clinical trials in T1D prevention studies. The following describes the methodology of the proposed tool.

A. The user will define a set of subject characteristics at study entry accounting for relevant covariates. In particular, the user will specify the relative proportions of

combinations of islet AAs, a threshold for 120-minute timepoint OGTT, and a threshold for HbA1c percentage.

B. A virtual patient population of N individuals will be generated with the characteristics defined in (A.)
C. The tool will use the underlying model to compute the survival curve for each of the N individuals. The curves will be averaged to determine the mean survival for the population defined by the chosen covariates.
D. Resampling techniques will be used to generate confidence intervals around the mean survival curve.
E. Ranges of predicted times to diagnosis for the chosen population will be computed using the mean survival curve and corresponding confidence intervals from D.
F. Based on results obtained in (E.), the user may decide to change entry characteristics (A.) and re-run the survival analysis (D.) until the desired patient population characteristics are achieved.

4.5.2 Key Deliverables

This briefing dossier is accompanied by R scripts and data files. The R scripts consists of 4 R markdown files, "1 Analysis subset derivation.Rmd", "2 Data analysis.Rmd", "3 Modeling analysis.Rmd" and "4 Model validation.Rmd". The data files consist of 3 Comma-Separated Values (CSV) files, "derived_bl_data_clean.csv" "longitudinal_derived_baseline.csv" and "daisydatamart_updated.csv". R scripts and data files are provided in a single zip file with an appropriate folder structure. The Appendix I provides instructions for running the R markdown files and viewing the results.

4.6 Conclusions: Methodology and Results

The goal of this work is to leverage existing data sources that captured islet AA measurements and glycemic markers in a population likely to participate in T1D prevention trials to generate evidence supporting a qualification opinion for the use of islet AAs as enrichment biomarkers for T1D prevention trials . The underlying evidence is presented as a time-to-event model for predicting the probability of T1D diagnosis. Data sources included multiple observational studies and required harmonization efforts to ensure interoperability. The resulting analysis set was used to test various models beginning, using a Cox PH model as an initial approach. The analysis for the Cox PH model started by verifying the validity of the PH assumption. As this assumption was not met and given the SAWP's recommendation to evaluate a parametric modeling approach, a parametric AFT model was chosen. Model diagnostic and performance exercises were performed to assess the model and quantify the effect of islet AAs, baseline age, sex, and glycemic markers on time to T1D diagnosis. The AFT model constituted the candidate model presented in this briefing dossier.

The use of baseline islet AA positivity in the model was represented as a single covariate with eleven distinct levels representing all possible combinations of two or more of GAD65, IAA, IA-2, and ZnT8. This approach is more granular than considering the total numbers of islet AAs and provides the ability to quantify risk by islet AA type. By only assessing individuals at baseline, it is possible that subjects positive for two or three islet AA may convert to the three or four islet AAs before diagnosis. However, this baseline selection method reflects how sponsors will recruit subjects for T1D prevention studies as the islet AA time history of subjects will not be available to sponsors. The use of baseline information is therefore preferred in this context.

Results from the AFT model indicate that GAD65_IAA and GAD65_ZnT8 combinations have the least relative risk compared to all other combinations, while the IA-2_ZnT8 has the highest relative risk. The presence of all four islet AAs has a marginal increase of risk relative to the

baseline hazard. The use of 120-minute OGTT, baseline age, sex and HbA1c values provide a significant ability to further stratify the risk of T1D diagnosis within these islet AA positive populations (alt_mod3). To provide credibility to these results, both internal and external validation procedures were carried out for selected model (alt_mod3). Internally, a time dependent receiver operation characteristic (ROC) analysis was performed, which showed high overall concordance across AUC values (> 0.75), especially within the first two-years following the derived baseline, which represents a time frame concordant with feasible trial design for T1D prevention. The internal validation was carried out to measure concordance in a time-dependent manner using k-fold cross validation, as model accuracy over the reasonable duration of prevention trials is the more important consideration. In this case, concordance in the first two years was high (c-index > 0.75). Additional internal validation was carried out for a pediatric population (<12 years); as such, this population is of keen interest to sponsors. Results showed a high degree of concordance in this population as well (c-index≈0.8). External validation was carried out on the DAISY study, additionally showing a high concordance in the first two years (c-index > 0.8), suggesting that the presented underlying evidence is adequate for a qualification opinion for the use of islet AAs for trial enrichment, as per the proposed COU.

A key consideration of the data is the selection of individuals with non-missing information for glycemic measurements for inclusion in the derived baseline population used in the AFT model. The applicant considers this population is representative of those likely to enter a T1D prevention study. The quantification of OGTT values on timing to T1D diagnosis are especially important for further stratification. Another consideration of the data is that the time history of islet AA positivity is unknown 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 is to qualify islet AAs as enrichment biomarkers for T1D prevention studies. As such, the applicant considers that a sponsor of a prevention T1D study will not know the islet AA time history of participating subjects, and the TN01 data discussed in this dossier are representative of a population that is likely to enter a T1D prevention study. The proposed use of islet AAs as enrichment tools for T1D prevention trials provides a basis to identify subpopulations likely to reach a T1D diagnosis during trials of reasonable duration. As such, a quantitative tool to quantify the variability of expected times to diagnosis is out of scope for this qualification opinion. Lastly, the size of the external validation was small due to the definition of the baseline used in the analysis. While the model performed well, additional credibility can be established with larger numbers of subjects from other independent datasets.

In conclusion, the applicant demonstrated that analysis of integrated data from independent observational data sources represents adequate supporting evidence for a qualification opinion 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. The adoption of this tool is expected to help stimulate drug development in T1D prevention.

5 SUMMARY AND CONCLUSIONS

Currently, the development of therapies to prevent or delay the onset of T1D remains challenging, and there is a lack of qualified biomarkers to identify individuals at risk of developing T1D or to quantify the risk of conversion to a diagnosis of T1D. There have been significant late-stage failures in the development of therapies in new-onset T1D. In order to address this drug development need, the T1DC has 1) acquired, remapped, integrated, and

curated existing patient-level data from observational studies and 2) evaluated the utility of islet AAs, including IAA, GAD65, IA-2, and ZnT8, as biomarkers to enrich subjects for T1D prevention trials using a model-based approach. The TEDDY and TN01 studies were aggregated to support the model-based qualification of islet AAs as enrichment biomarkers. This aggregated dataset was used to construct and execute a statistical analysis plan to develop a time-to-event model for predicting T1D diagnosis. The developed model demonstrates that islet AAs are statistically significant predictors of the time-varying probability of conversion to a diagnosis of T1D, representing adequate underlying evidence for their use as enrichment tools. Further when additional sources of variability, including baseline age, sex, blood glucose measurements from the 120-minute timepoints of OGTT, and HbA1c, are assessed with the islet AAs, it improves the accuracy of predicting the time-varying probability of conversion to a T1D diagnosis.

The proposed COU focuses on the application of islet AAs with other patient features as enrichment biomarkers to optimize the selection of subjects for clinical trials of therapies intended to prevent or delay the clinical diagnosis of T1D. The focus of this briefing dossier is to present the underlying evidence to support this qualification opinion. From a practical drug development perspective, the use of islet AAs as enrichment biomarkers can improve clinical trial design by informing trial entry criteria, improving stratification approaches, and adding more efficiency into the drug development process for critical therapies that may prevent or delay T1D. T1DC gratefully acknowledges the EMA for the public posting of the EMA Letter of Support, "*Islet autoantibodies as enrichment biomarkers for type 1 diabetes prevention studies, through a quantitative disease progression model*", posted on 25 March 2020. A full qualification opinion will serve to encourage widespread use of the proposed quantitative tool for enrichment strategies and stratification in ongoing and future clinical trials.

6 QUESTIONS FOR EMA FOLLOWED BY T1DC's POSITION

1. **Does EMA agree with the COU?**

   T1DC's position: The proposed COU focuses on the application of islet AAs, together with other patient features, as enrichment biomarkers in individuals at risk of developing T1D to optimize the selection of individuals for clinical trials of therapies intended to prevent or delay the clinical diagnosis of T1D. The focus is on understanding the contribution of the positivity to these AAs as predictors of progressing towards a diagnosis of T1D. From a practical drug development standpoint, this proposed use is of added value because their intended application can help inform the definition of entry criteria, enrichment strategies, and stratification approaches in the field of T1D prevention.

2. **Does EMA agree that the data sources are adequate to support the proposed COU?**

   T1DC's position: The available data sources, and their integration through data standardization and management, represents a unique opportunity to transform these data into valuable knowledge to provide the necessary evidence to support the qualification of islet AAs for the proposed context of use. The population captured in the data sources represents the population likely to be considered as candidates to participate in clinical trials of therapies intended to prevent or delay the clinical diagnosis of T1D.

3. **Does EMA agree the AFT survival model and its covariates represent adequate evidence for the qualification of islet AAs as enrichment biomarkers for T1D prevention trials?**

   T1DC's position: The T1DC believes a survival model construct is adequate because the clinically relevant endpoint defined for the proposed model is a binary dependent variable and the need to understand the likelihood of conversion to a diagnosis of T1D over the course of clinical for prevention or delay of T1D. The proposed survival model evaluating the contribution of subject's positivity to the different islet AAs taken in combination to understand the time-varying probability of conversion to a diagnosis of T1D also represents an adequate approach to provide the supporting evidence for this intended qualification procedure.

4. **Does EMA agree that the validation is adequate?**

   T1DC's position: The k-fold cross-validation approach is an adequate method to assess model performance, given all observations are used for training and validation and each observation is used for validation exactly once. This approach has been successfully used in prior qualification procedures with EMA for different novel methodologies in drug development, including biomarkers and quantitative drug development tools. While additional validation using published meta-data was not deemed feasible, an additional external independent patient-level dataset, (i.e., DAISY), was acquired by the T1DC and used to perform patient-level external validation. This approach provided further evidence of robust model performance.

5. **Does EMA agree the presented results represent adequate supporting evidence for a qualification opinion?**

   T1D Consortium position: The presented results demonstrate that the combinations of islet AA for which subjects are seropositive at a sensible baseline for clinical trials

independent and statistically significant time-varying predictors of T1D. The presented analyses also show that the use of positivity for combinations of islet AAs together with measures of glycemic control can help inform the definition of entry criteria, enrichment strategies, and stratification approaches for T1D prevention clinical trials.

7 REFERENCES

"ASK Research Program / Autoimmunity Screening for Kids / Denver, CO." n.d. Ask-Health. Accessed April 16, 2019. https://www.askhealth.org.

Barkai, László, Zoltán Kiss, György Rokszin, Zsolt Abonyi-Tóth, György Jermendy, István Wittmann, and Péter Kempler. 2020. "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." *Archives of Medical Science: AMS* 16 (1): 34–41. https://doi.org/10.5114/aoms.2019.88406.

Battaglia, Manuela, Mark S. Anderson, Jane H. Buckner, Susan M. Geyer, Peter A. Gottlieb, Thomas W. H. Kay, Åke Lernmark, et al. 2017. "Understanding and Preventing Type 1 Diabetes through the Unique Working Model of TrialNet." *Diabetologia* 60 (11): 2139–47. https://doi.org/10.1007/s00125-017-4384-2.

Breiman, Leo, and Philip Spector. 1992. "Submodel Selection and Evaluation in Regression. The X-Random Case." *International Statistical Review / Revue Internationale de Statistique* 60 (3): 291. https://doi.org/10.2307/1403680.

Burnham, Kenneth P., and David R. Anderson. 2016. "Multimodel Inference: Understanding AIC and BIC in Model Selection." *Sociological Methods & Research*, June. https://doi.org/10.1177/0049124104268644.

Foster, Nicole C., Roy W. Beck, Kellee M. Miller, Mark A. Clements, Michael R. Rickels, Linda A. DiMeglio, David M. Maahs, et al. 2019. "State of Type 1 Diabetes Management and Outcomes from the T1D Exchange in 2016-2018." *Diabetes Technology & Therapeutics* 21 (2): 66–72. https://doi.org/10.1089/dia.2018.0384.

Hagopian, William A, Henry Erlich, Åke Lernmark, Marian Rewers, Anette G Ziegler, Olli Simell, Beena Akolkar, et al. 2011. "The Environmental Determinants of Diabetes in the Young (TEDDY): Genetic Criteria and International Diabetes Risk Screening of 421 000 Infants." *Pediatric Diabetes* 12 (8): 733–43. https://doi.org/10.1111/j.1399-5448.2011.00774.x.

Harrell, Frank E. 2019. *Rms: Regression Modeling Strategies* (version 5.1-4). https://CRAN.R-project.org/package=rms.

Heagerty, PJ, and Y Zheng. 2005. "Survival Model Predictive Accuracy and ROC Curves. - PubMed - NCBI." 2005. https://www.ncbi.nlm.nih.gov/pubmed/15737082.

Insel, Richard A., Jessica L. Dunne, Mark A. Atkinson, Jane L. Chiang, Dana Dabelea, Peter A. Gottlieb, Carla J. Greenbaum, et al. 2015. "Staging Presymptomatic Type 1 Diabetes: A Scientific Statement of JDRF, the Endocrine Society, and the American Diabetes Association." *Diabetes Care* 38 (10): 1964–74. https://doi.org/10.2337/dc15-1419.

International Diabetes Federation. 2019. "IDF Atlas 9th Edition and Other Resources." IDF Diabetes Atlas, 9th Edn. 2019. https://www.diabetesatlas.org.

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.

Kassambara, Alboukadel, and Marcin Kosinski. 2018. "Survminer: Drawing Survival Curves Using 'Ggplot2'." *URL Https://CRAN. R-Project. Org/Package= Survminer. R Package Version 0.4* 3: 15.

Kick, Kerstin, Verena S. Hoffmann, Karin Lange, Martin Lang, Otto Laub, Susanne Bechtold-Dalla Pozza, Dominik Böcker, et al. 2018. "Feasibility and Organization of a Population-Based Screening for Pre-Symptomatic Type 1 Diabetes in Children — Evaluation of the Fr1da Study." *Journal of Public Health*, September. https://doi.org/10.1007/s10389-018-0981-x.

Köhler, Meike, Andreas Beyerlein, Kendra Vehik, Sonja Greven, Nikolaus Umlauf, Åke Lernmark, William A. Hagopian, et al. 2017. "Joint Modeling of Longitudinal Autoantibody Patterns and Progression to Type 1 Diabetes: Results from the TEDDY

Study." *Acta Diabetologica* 54 (11): 1009–17. https://doi.org/10.1007/s00592-017-1033-7.
Köhler, Meike, Nikolaus Umlauf, Andreas Beyerlein, Christiane Winkler, Anette-Gabriele Ziegler, and Sonja Greven. 2017. "Flexible Bayesian Additive Joint Models with an Application to Type 1 Diabetes Research." *Biometrical Journal* 59 (6): 1144–65. https://doi.org/10.1002/bimj.201600224.
Mahon, Jeffrey L., Jay M. Sosenko, Lisa Rafkin-Mervis, Heidi Krause-Steinrauf, John M. Lachin, Clinton Thompson, Polly J. Bingley, et al. 2009. "The TrialNet Natural History Study of the Development of Type 1 Diabetes: Objectives, Design, and Initial Results." *Pediatric Diabetes* 10 (2): 97–104. https://doi.org/10.1111/j.1399-5448.2008.00464.x.
Ozenne, Brice, Anne Sørensen Lyngholm, Thomas Scheike, Christian Torp-Pedersen, and Thomas Gerds Alexander. 2017. "RiskRegression: Predicting the Risk of an Event Using Cox Regression Models." *The R Journal* 9 (2): 440. https://doi.org/10.32614/RJ-2017-062.
Potapov, Sergej, Werner Adler, and Matthias Schmid. 2015. *Package 'SurvAUC', Estimators of Prediction Accuracy for Time-to-Event Data.* (version 1.0-5).
Raab, Jennifer, Florian Haupt, Marlon Scholz, Claudia Matzke, Katharina Warncke, Karin Lange, Robin Assfalg, et al. 2016. "Capillary Blood Islet Autoantibody Screening for Identifying Pre-Type 1 Diabetes in the General Population: Design and Initial Results of the Fr1da Study." *BMJ Open* 6 (5). https://doi.org/10.1136/bmjopen-2016-011144.
Rewers, M., T. L. Bugawan, J. M. Norris, A. Blair, B. Beaty, M. Hoffman, R. S. McDuffie, et al. 1996. "Newborn Screening for HLA Markers Associated with IDDM: Diabetes Autoimmunity Study in the Young (DAISY)." *Diabetologia* 39 (7): 807–12.
Rewers, M., J. M. Norris, G. S. Eisenbarth, H. A. Erlich, B. Beaty, G. Klingensmith, M. Hoffman, et al. 1996. "Beta-Cell Autoantibodies in Infants and Toddlers without IDDM Relatives: Diabetes Autoimmunity Study in the Young (DAISY)." *Journal of Autoimmunity* 9 (3): 405–10.
Skrivarhaug, Torild, Lars C. Stene, Ann Kristin Drivvoll, Hanne Strøm, Geir Joner, and Norwegian Childhood Diabetes Study Group. 2014. "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 (1): 57–62. https://doi.org/10.1007/s00125-013-3090-y.
Skyler, Jay S., Jeffrey P. Krischer, Joseph Wolfsdorf, Catherine Cowie, Jerry P. Palmer, Carla Greenbaum, David Cuthbertson, Lisa E. Rafkin-Mervis, H. Peter Chase, and Ellen Leschek. 2005. "Effects of Oral Insulin in Relatives of Patients with Type 1 Diabetes: The Diabetes Prevention Trial--Type 1." *Diabetes Care* 28 (5): 1068–76.
Sosenko, Jay M., Jerry P. Palmer, Lisa Rafkin-Mervis, Jeffrey P. Krischer, David Cuthbertson, Jeffery Mahon, Carla J. Greenbaum, Catherine C. Cowie, Jay S. Skyler, and Diabetes Prevention Trial-Type 1 Study Group. 2009. "Incident Dysglycemia and Progression to Type 1 Diabetes among Participants in the Diabetes Prevention Trial-Type 1." *Diabetes Care* 32 (9): 1603–7. https://doi.org/10.2337/dc08-2140.
Steck, Andrea K., Kendra Vehik, Ezio Bonifacio, Ake Lernmark, Anette-G. Ziegler, William A. Hagopian, JinXiong She, et al. 2015. "Predictors of Progression From the Appearance of Islet Autoantibodies to Early Childhood Diabetes: The Environmental Determinants of Diabetes in the Young (TEDDY)." *Diabetes Care* 38 (5): 808–13. https://doi.org/10.2337/dc14-2426.
Szalecki, Mieczysław, Marta Wysocka-Mincewicz, Anna Ramotowska, Artur Mazur, Lucyna Lisowicz, Iwona Beń-Skowronek, Joanna Sieniawska, et al. 2018. "Epidemiology of Type 1 Diabetes in Polish Children: A Multicentre Cohort Study." *Diabetes/Metabolism Research and Reviews* 34 (2). https://doi.org/10.1002/dmrr.2962.
Therneau, Terry M. 2020. "Package 'Survival' Version 3.1-12." https://cran.r-project.org/web/packages/survival/survival.pdf.

Veijola, Riitta, Maarit Koskinen, Olli Helminen, and Anne Hekkala. 2016. "Dysregulation of Glucose Metabolism in Preclinical Type 1 Diabetes." *Pediatric Diabetes* 17 Suppl 22: 25–30. https://doi.org/10.1111/pedi.12392.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and RStudio. 2020. *Dplyr: A Grammar of Data Manipulation* (version 0.8.5). https://CRAN.R-project.org/package=dplyr.

Writing Group for the DCCT/EDIC Research Group, Trevor J. Orchard, David M. Nathan, Bernard Zinman, Patricia Cleary, David Brillon, Jye-Yu C. Backlund, and John M. Lachin. 2015. "Association between 7 Years of Intensive Treatment of Type 1 Diabetes and Long-Term Mortality." *JAMA* 313 (1): 45–53. https://doi.org/10.1001/jama.2014.16107.

Xu, Ping, Jeffrey P. Krischer, and Type 1 Diabetes TrialNet Study Group. 2016. "Prognostic Classification Factors Associated With Development of Multiple Autoantibodies, Dysglycemia, and Type 1 Diabetes-A Recursive Partitioning Analysis." *Diabetes Care* 39 (6): 1036–44. https://doi.org/10.2337/dc15-2292.

Yoshida, Kenta, and Laurent Claret. 2020. *SurvParamSim: Parametric Survival Simulation with Parameter Uncertainty* (version 0.1.2). https://CRAN.R-project.org/package=survParamSim.

Ziegler, AG, Kerstin Kick, Ezio Bonifacio, Florian Haupt, Markus Hippich, Desiree Dunstheimer, Martin Lang, et al. 2020. "Yield of a Public Health Screening of Children for Islet Autoantibodies in Bavaria, Germany." *JAMA* 323 (4): 339–51. https://doi.org/10.1001/jama.2019.21565.

Ziegler, Anette G., Marian Rewers, Olli Simell, Tuula Simell, Johanna Lempainen, Andrea Steck, Christiane Winkler, et al. 2013. "Seroconversion to Multiple Islet Autoantibodies and Risk of Progression to Diabetes in Children." *JAMA* 309 (23): 2473. https://doi.org/10.1001/jama.2013.6285.