

28 March 2022 EMA/CHMP/SAWP/186420/2022 Committee for Medicinal Products for Human Use (CHMP)

Qualification Opinion of Islet Autoantibodies (AAs) as Enrichment Biomarkers for Type 1 Diabetes (T1D) Prevention Clinical Trials

| Draft agreed by Scientific Advice Working Party (SAWP) | 11 Feb 2021 |
|--|-------------------------------|
| Adopted by CHMP for release for consultation | 25 March 2021 ¹ |
| Start of public consultation | 03 November 2021 2 |
| End of consultation (deadline for comments) | 14 December 2021 ³ |
| Adoption by CHMP | 24 March 2022 |

| Keywords | Qualification of Novel Methodology, enrichment marker, patient selection, |
|----------|--|
| | Type 1 Diabetes Mellitus, studies investigating prevention or delay in disease |
| | onset, islet autoantibodies, anti-IAA, anti-GAD65, anti-IA-2, anti-ZnT8, |
| | accelerated time-failure model |



¹ Last day of relevant Committee meeting.

² Date of publication on the EMA public website.

³ Last day of the month concerned.

1. Executive summary

The objective of this procedure was 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. The proposed context-of-use (COU) statement was that, 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 an oral glucose tolerance test (OGTT) and haemoglobin A1c (HbA1c) levels.

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)⁴. 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. An accelerated time failure model provides 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 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, sex, baseline age, 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 timevarying 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.

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 sex, baseline age, blood glucose measurements from the 120-minute timepoints of an OGTT and HbA1c. The process by which these sources of variability were selected is outlined.

The developed models were shown to demonstrate that the baseline presence of various combinations 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, sex, and baseline age 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.

2. Answers to applicant's questions

Based on the coordinators' reports the CHMP gave the following answers to the questions by the applicant:

Question 1:

Does EMA agree with the COU?

⁴The data from the TEDDY and TrialNet Study reported here were supplied by the NIDDK Central Repositories. This document/publication does not necessarily reflect the opinions or views of the TEDDY, TrialNet Study, the NIDDK Central Repositories, or the NIDDK.

Qualification Opinion of Islet Autoantibodies (AAs) as Enrichment Biomarkers for Type 1 Diabetes (T1D) Prevention Clinical Trials EMA/CHMP/SAWP/186420/2022

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.

CHMP answer

The qualification exercise included a modeling exercise that also identified the relevance of additional clinical parameters (sex, baseline age, blood glucose measurements from the 120-minute timepoints of oral glucose tolerance test (OGTT), and haemoglobin A1c (HbA1c) levels).

Individuals defined as 'At risk' were defined in this context as being a first degree relative (FDR) of a T1D patient or those having a specific human leukocyte antigen (HLA) subtype of risk (HLA-DR3/3, DR4/4, DR3/4, DR3/X [X \neq 3], DR4/X [X \neq 4]), excluding individuals with baseline fasting glucose \geq 126mg/dL (7.0 mmol/L) or stimulated 2-hour glucose \geq 200 mg/dL (11.1 mmol/L).

Positivity for two or more of the islet AAs, determined in this population, in addition to the relevant characteristics as described in the model, can be used for enrichment of clinical trials focusing on the delay or prevention of the clinical diagnosis of T1D.

The proposed COU is overall agreed. The clinical interest of identifying good biomarkers for Type 1 Diabetes (T1D) onset in an at-risk patient population is supported by the CHMP, and the unmet need for better means to optimize drug development in the field is acknowledged. There are no approved therapies to prevent or delay the onset of T1D and there is a lack of biomarkers to identify individuals and quantify risk of conversion to a diagnosis of T1D. In addition, there have been significant failures in late-stage development of therapies in new-onset T1D. These failures have been attributed in part to a high degree of heterogeneity in the patient population and a current inability to quantitatively describe the contributions of specific sources of variability to such heterogeneity. Second, intervening in new-onset T1D may be too late to significantly delay or halt disease progression and preserve endogenous β -cell function.

A practical problem foreseen is that in clinical trial recruitment, often the only parameter known is family history, which could limit the utility of this new screening/enriching tool unless mass screening efforts are taking place. During the discussion meeting (DM), the applicant clarified that there would also be the possibility to partner with pre-existing trial networks such as TrialNet and INNODIA that are carrying out screening efforts. This is supported.

The model-based approach proposed by the applicant is considered an acceptable method to address the question of interest, which is whether the combination of positivity to 2 or more of the 4 selected AAs can be considered acceptable predictors of a diagnosis of T1D, when combined with additional and well-defined patient characteristics.

It should however be noted that the modeling approach taken by the applicant is not a mechanistic disease model: a clear and fully quantitative description of the contribution of the different factors including positivity to these AAs as predictors of progressing towards a diagnosis of T1D is therefore not possible. The model allows confirming the existence of a significant statistical contribution of the different covariates and their relative relevance toward T1D diagnosis for patients at risk.

The analytical assays used to measure islet autoantibodies (AA) against glutamic acid decarboxylase 65 (GAD65), insulinoma antigen-2 (IA-2; also called ICA512), insulin (IAA) and zinc transporter 8 AA (ZnT8) are considered state of the art. It should be noted that the results and the conclusions of the modeling analysis as assessed during this qualification procedure are considered only applicable when

the islet autoantibodies are measured using these methods or methods proved to have at least equivalent analytical performances.

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 \neq 3], DR4/X [X \neq 4]), excluding individuals with baseline fasting glucose \geq 126mg/dL (7.0 mmol/L) or stimulated 120-minute glucose \geq 200 mg/dL (11.1 mmol/L). 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 for 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 sex, baseline age, baseline HbA1c levels and the 120-minute time point 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. The underlying time-to-event models that supports this qualification will be made available through the Critical Path Institute's website (https://www.c-path.org/).

Out-of-scope:

• The underlying evidence for the COU does not account for variability in the longitudinal seroconversion for the different islet AAs over the course of T1D prevention trials.

• The underlying time-to-event model that provides the evidence for the COU statement of the qualified biomarkers does not include the ability to generate virtual sub-populations for simulation purposes.

The COU is overall agreed. There is clearly an unmet need for biomarkers to aid development in T1DM prevention, a field with a long history of failed trials. This Qualification would only refer to the value of the positivity of two AAs in the risk assessment. The combination of AAs, numbers above two AAs and the reason for not assessing only one AA are well explained. With a validated method, this would clearly help with selection and stratification of subjects in clinical development. Having a model of the effect of two positive AAs cannot replace a placebo arm in a randomized trial setting.

Question 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.

CHMP answer

The data used for the model development and external validations to support the qualification of islet AAs as enrichment biomarkers originated from three datasets: the TN01, TEDDY, and DAISY registry studies. A summary of the three studies can be found in Table 1.TEDDY and TN01 were aggregated and used for model development and internal cross-validation. Data from the DAISY study was acquired and used to perform external validation on the final model.

Participants for TN01 were selected by the presence of a FDR with T1D, as this has been shown to be a risk factor for 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 1 islet AA were considered eligible for follow-up. In 2009 the eligibility for follow-up changed to the presence of 2 islet AAs. As of December 2018, 4,524 subjects are being followed. Once subjects were selected for follow-up and opted in, they were monitored for 6 monthly visits using oral glucose tolerance test (OGTT), detection of islet AAs and measurement of HbA1c levels. TN01 is providing TrialNet with an active patient ready cohort and collaborative clinical trial network to evaluate novel therapies. TN01 is still enrolling new subjects and following current subjects. The TN01 data provided in this submission is locked as of December 2018.

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. Children were screened and recruited during infancy based on high-risk HLA genotypes (361,518 initial screenings and 8,667 in initial prospective cohort), with separate inclusion criteria for GP children or FDR. Participants are monitored prospectively with study visits every 3 months for the first 4 years, and every 6 months thereafter to age 18. All children who are persistently positive for any islet AA are monitored every 3 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 participants, 89% had no family history of T1D.

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 et al. 1996a; Rewers et al. 1996b). 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. 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 both were based on the American Diabetes Association (ADA) criteria. All studies are observational but certain features in their designs differ, including inclusion criteria and scheduled frequency of follow-up.

The data sources are judged largely relevant, consistent with the recommendation during the QA procedure. From a modeling perspective, this approach is endorsed, and the 3 data sources seem adequate. Potential covariate distribution and correlation were presented and discussed as requested during the qualification procedure.

The baseline data intended for modeling are relatively well defined, as well as the binary endpoint (T1D diagnosis).

Qualification Opinion of Islet Autoantibodies (AAs) as Enrichment Biomarkers for Type 1 Diabetes (T1D) Prevention Clinical Trials EMA/CHMP/SAWP/186420/2022

Longitudinal assessments of islet AA positivity, OGTTs, C-peptide measurements, and HbA1c measurements are considered out of scope for the proposed analysis, and only baseline information were used for the modeling analysis.

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

• Presence of any two or more of the 4 islet AAs

• Complete, (i.e., non-missing) information for OGTT (0 and 120-minute time points), C-peptide measurements (0 and 120-minute time points), HbA1C measurements, age and sex.

1 Table 1. Overview TN01, TEDDY, and DAISY

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

2 * FDR is defined as a child, parent, or sibling.

3 ** Extended family member is defined as a cousin, niece, nephew, aunt, uncle, grandparent, or half-sibling.

- 4 *** 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,
- 5 DR3/4, DR3/X [X≠3], DR4/X [X≠4]
- 6 + Between 2004-2009 individuals with one islet AA were followed with six-monthly assessments. After 2009 this changed, and subjects required two or more
- 7 islet AAs to be enrolled in the follow-up cohort

8 **Question 3:**

9 Does EMA agree the AFT survival model and its covariates represent adequate evidence for

10 the qualification of islet AAs as enrichment biomarkers for T1D prevention trials?

11 T1DC's position: T1DC believes a survival model construct is adequate because the clinically relevant

12 endpoint defined for the proposed model is a binary dependent variable and the need to understand

13 the likelihood of conversion to a diagnosis of T1D over the course of a clinical trial for prevention or

delay of T1D. The proposed survival model evaluating the contribution of subject's positivity to the 14

- 15 different islet AAs taken in combination to understand the time-varying probability of conversion to a
- 16 diagnosis of T1D also represents an adequate approach to provide the supporting evidence for this
- 17 intended qualification procedure.

18 **CHMP** answer

19 The applicant developed a survival model to describe the time course of incidence of T1DM in patients

20 included in the 2 datasets used for model building (TEDDY and TN01), given their baseline

characteristics. The third dataset was used for model validation. The following hazard functions were 21

tested and compared based on their Akaiké information criteria during the modeling process: Weibull, 22

- 23 gamma, generalized gamma, generalized F, log logistic distributions. The patient baseline
- 24 characteristics tested as covariates in the model, as well as their brief description are included in table
- 25 3 below. Table 4 and 5 provide their respective descriptive statistics.

| Notation | Description of covariate at derived | Туре |
|--|--|------------|
| | baseline | |
| X_{GAD65_IAA} | Positivity for GAD65, IAA | Binary |
| X_{GAD65_IA-2} | Positivity for GAD65, IA-2 | Binary |
| X _{GAD65_ZnT8} | Positivity for GAD65, ZnT8 | Binary |
| $X_{IA-2_{IAA}}$ | Positivity for IA-2, IAA | Binary |
| X_{IA-2_ZnT8} | Positivity for IA-2, ZnT8 | Binary |
| X _{IAA_ZnT8} | Positivity for IAA, ZnT8 | Binary |
| X _{GAD65_IAA_ZnT8} | Positivity for GAD65, IAA, ZnT8 | Binary |
| $X_{GAD65_IAA_IA-2}$ | Positivity for GAD65, IAA, IA-2 | Binary |
| $X_{\text{GAD65}_{\text{IA}-2}_{\text{ZnT8}}}$ | Positivity for GAD65, IA-2, ZnT8 | Binary |
| $X_{IA-2_IAA_ZNT8}$ | Positivity for IA-2, IAA, ZnT8 | Binary |
| $X_{GAD65_IA-2_IAA_ZNT8}$ | 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- | Continuous |
| | minute results from OGTT | |
| X _{Log_GLU120_s} | Log transformed and standardized and | Continuous |
| | 120-minute results from OGTT | |

Table 2 Covariates evaluated 26

27 28 29 * High-risk HLA is defined in <u>Section 4.3.3.2</u>

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

30

31 Table 3. 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 |

32 **Table 4. T1D diagnoses in the analysis set by autoantibody combination**

| | Т | EDDY | TN01 | | | |
|-----------------|----------|-----------|--------------|----------|-----------|--------------|
| Islet AA | Subjects | Diagnoses | % Conversion | Subjects | Diagnoses | % Conversion |
| combination | | | | | | |
| GAD65_IA-2 | 34 | 15 | 44% | 150 | 35 | 23% |
| GAD65_IA-2_IAA | 28 | 13 | 46% | 64 | 16 | 25% |
| GAD65_IA- | 74 | 39 | 53% | 280 | 83 | 30% |
| 2_IAA_ZnT8 | | | | | | |
| 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% |

³³ Given the empirical nature of the model, the results obtained by the applicant are also considered

34 highly dependent on tested covariate distribution and correlation/collinearity.

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

- 41 The correlation between the continuous covariates (Figure 4) did not reveal any covariate pairs with
- 42 high correlation, defined as correlations above 0.3. The Wilcoxon test (Table 11) and the chi-square
- 43 test of independence (Table 12) showed that the baseline Age (bAGE_s) and SEX were highly
- 44 associated with AA combinations. Association between islet AA combinations was not considered
- 45 relevant as their presence is mutually exclusive (i.e., only one islet AA combination is possible for a
- 46 given subject at a single measurement).

47

Qualification Opinion of Islet Autoantibodies (AAs) as Enrichment Biomarkers for Type 1 Diabetes (T1D) Prevention Clinical Trials EMA/580542/2021

48 *Figure 4. Pearson's correlation between continuous covariates*



Table 5. 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 |

51 Table 6. 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 |

52 Modeling Analysis Methodology

| 53 As per the original statistical analysis plan, the first approach was to analyze predic | ors of T1D: |
|--|-------------|
|--|-------------|

54 diagnosis using a Cox proportional hazard (PH) model, (i.e., a semi-parametric approach), as this was

55 the most parsimonious first step. Based on reviewer recommendations, a fully parametric approach

56 was requested. With knowledge of prior quantitative analyses from the literature, consideration of the

57 drug development context, and the available data, the full modeling analysis was executed. The flow

- 57 drug development context, and the available data, the full modeling analysis was executed. The not
- 58 chart (Figure 3) displays the progression of the modeling analysis, where subsequent steps were

- 59 executed based on best practices for model building and learnings from previous steps. All analysis
- 60 was carried out in the R programming language. In completion, the model building process followed
- 61 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
- 63 TEDDY datasets; (c) Evaluation of model performance with k-fold cross-validation and external
- 64 validation with DAISY as a separate independent dataset.



65 Software

- 66 Model building, visualization, model assumptions, diagnostics and external validation was conducted in
- 67 R (version 4.0.0; Vienna, Austria, R Core Team, 2018) using the packages "survival" (Therneau 2020),
- 68 "flexsurv" (Jackson 2016), "survminer" (Kassambara and Kosinski, n.d.), "dplyr" (Wickham et al.
- 69 2020), "survAUC" (Potapov, Adler, and Schmid 2015), "rms" (Harrell 2019) and "riskRegression"

 $h_{i}(t) = h_{0}(t) \exp(\sum_{i \in I} \beta_{i} X_{ii})$

70 (Ozenne et al. 2017).

71 Cox Proportional Hazard Model

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

- 74 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

(E1)

- vise 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
- 78 individuals remains constant over time).
- 79 Selection of Parametric Distribution

- 80 Multiple parametric distributions were tested for their ability to approximate the underlying hazard
- 81 function including exponential, Weibull, gamma, generalized gamma, generalized F, log logistic, log
- 82 normal and Gompertz. Resulting Akaike information criterion (AIC) values and graphical methods for
- 83 survival and hazard function fits were compared to select an appropriate parametric form. The
- 84 'flexsurvreg' function in the 'flexsurv' R package was used for the selection of parametric distribution
- 85 analysis.
- 86 Univariate Analysis
- A univariate analysis was performed by estimating a Cox PH model for of the covariates in Table 3. The
- 88 'coxph' function in the 'survival' R package was used for Cox PH analysis (Therneau 2020). Covariates
- 89 with no significant univariate association (p-value \geq 0.1) with T1D diagnosis were not considered for
- 90 the full model development. The p-value was computed using the Wald test, which evaluates whether
- 91 the covariate coefficient is statistically different from zero. A multiplicity adjusted alpha value
- 92 (Bonferroni correction) was used for univariate analysis.
- 93 Analysis of Correlation and Association between Covariates
- 94 The covariates remaining after the univariate analysis were analyzed for multicollinearity and
- 95 associations prior to performing multivariate analysis. Pearson's correlation was used to test the
- 96 correlation between continuous covariates, with a correlation value above 0.3 chosen as significant.
- 97 The Wilcoxon test was used to test the association between continuous and categorical covariates, and
- 98 the Chi-square test of independence was used to test the association between categorical covariates.
- 99 In both cases, a p-value < 0.001 (multiplicity adjusted) was chosen as the threshold for significance.
- 100 Multivariate Analysis
- 101 The multivariate analysis was performed by testing all possible combinations of remaining covariates,
- 102 as the number of covariates for multivariate analysis were reasonable. The comparison between
- 103 possible models was conducted using Akaike's Information Criteria (AIC). A reduction in AIC value
- 104 greater than or equal to 10 suggests a strong evidence in favor of the model with lower AIC (Burnham
- 105 and Anderson 2016).
- 106 *Model Diagnostics*
- 107 To assess if the PH assumption was satisfied, Schoenfeld residuals were utilized. The expected value of
- 108 these residuals can be used to quantify potential time-dependency on survival times. The Pearson
- 109 product-moment correlation between the scaled Schoenfeld residuals and log(time) for each covariate
- 110 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.

113 Parametric Accelerated Failure Time Model

- 114 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,
- 117

$$h_i(t) = h_0\left(t/\exp(\sum_{j \in I} \beta_j X_{ij})\right) \exp(-\sum_{j \in I} \beta X_{ij}) \quad (\mathsf{E2})$$

- 118 where $h_i(t)$ is hazard function for individual i determined by a set of j covariates $\{X_{ij}\}$ and
- 119 corresponding (estimated) coefficients $\{\beta_j\}$, t is the survival time, and $h_0(t)$ is the baseline hazard
- 120 defined by a parametric form with an underlying probability distribution such as Weibull, exponential,
- 121 or gamma. The β -parameter value specifies the effect each covariate has on the survival time, where
- 122 negative β values indicate that the survival time increases with positive-valued covariates, and positive
- 123 β values indicate that the survival time decreases with positive-valued covariates.
- 124 Selection of Parametric Distribution
- 125 Multiple parametric distributions were tested for their ability to approximate the underlying hazard
- 126 function including exponential, Weibull, gamma, generalized gamma, generalized F, log logistic, log
- 127 normal and Gompertz. Resulting Akaike information criterion (AIC) values and graphical methods for
- 128 survival and hazard function fits were compared to select an appropriate parametric form. The

- 129 'flexsurvreg' function in the 'flexsurv' R package was used for the selection of parametric distribution
- 130 analysis.
- 131 Univariate Analysis
- 132 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
- 134 'flexsurv' R package was used to perform parametric AFT model analysis. Individual covariates with no
- 135 significant association (P-value \geq 0.05) with T1D diagnosis were not considered for the full model
- 136 development. The p-value was computed using the Wald test, as described. 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.
- 139 Analysis of Correlation and Association between Covariates
- 140 The analysis defined in Section 4.3.5.3 was repeated for the covariates remaining after the AFT
- 141 univariate analysis.
- 142 Multivariate Analysis
- 143 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
- 146 greater than or equal to 10 suggests a strong evidence in favor of the model with lower AIC (Burnham
- 147 and Anderson 2016).
- 148 *Model Diagnostics*
- 149 Quantile-Quantile (Q-Q) plots were used to assess the validity of the AFT model assumption for two
- 150 groups of survival data. In this case, such groups correspond to the presence or absence of an AA
- 151 combination. Under the AFT model assumption, the presence of one islet AA combination has a
- 152 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
- 156 model. To analyze continuous covariates, binary groups were formed using thresholds to allow for the
- 157 generation of Q-Q plots.

158 Model Performance and internal Validation

- 159 Model Performance
- 160 To assess the model's predictive performance on the analysis set, time-dependent receiver operating
- 161 characteristic (ROC) curves were generated (Heagerty and Zheng 2005). Conceptually, the
- 162 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
- 164 computed. This is repeated across multiple timepoints to generate ROC curves. The area under the
- 165 ROC curves (AUC) are computed, which are interpreted as the concordance between the model
- 166 prediction and data. This methodology is an appropriate model performance metric as an individual's
- 167 risk for developing T1D changes over time. Further, it provides metrics as to the model's predictive
- 168 power for time frames over which a trial of reasonable duration would be conducted.
- 169 *K-fold cross validation*
- 170 Model validation was performed using the k-fold cross-validation technique (Breiman and Spector
- 171 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
- 175 model estimated survival on Kaplan-Meier curves for all five folds. The concordance index was
- 176 computed for each of the five folds estimated by time increments of one year up to six years.
- 177 Goodness-of-fit plots were created for visual assessments of models fits.
- 178 Cross-validation on Paediatric population

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

184 Model External Validation

- 185 External validation was performed using the DAISY dataset described. The definition of the derived
- 186 baseline was applied to the data to arrive at a validation set. The AFT model within this subset.
- 187 Goodness-of-fit plots were created by overlaying model estimated survival on Kaplan-Meier curves.
- 188 The concordance index was computed for time increments of one year up to six years.

189 *Modeling results*

190 A parametric AFT model was chosen using a Weibull distribution. Model diagnostic, performance, and

validation exercises were performed to assess the model's ability to quantify the time-varying effect of

islet AAs and glycaemic markers on risk to T1D diagnosis with overall satisfactory results. Results of

193 univariate and multivariate modeling are included in tables 17 and 19 below.

194 Table 7. 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 |

Qualification Opinion of Islet Autoantibodies (AAs) as Enrichment Biomarkers for Type 1 Diabetes (T1D) Prevention Clinical Trials EMA/580542/2021

195 Table 8. 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 |

196 The time-dependent ROC curves and AUC values showed good prediction performance, especially for

197 up to 2.5 years with AUC values greater than 0.8 (Figure 8).

198

Figure 8. Evaluation of model performance using time dependent Receiver operating characteristic (ROC) analysis on Final AFT model



201

202 Cross-Validation on Paediatric Population

203 The paediatric population (age < 12) in the analysis dataset comprised of 1330 subjects, with 345 204 from TEDDY and 985 from TN01. Half of this population i.e. 665 were randomly selected as test set for 205 this cross-validation analysis. A c-index of 0.8 or higher was obtained till 2 years and c-index of 206 0.75 or higher were obtained up to 6 years indicating good model performance (Table 20) 207 The visual predictive check (VPC) performed on the survival plot for cross-validation on the 208 paediatric population (age < 12) showed reasonable graphical fit (Figure 10). The dotted 209 curve represents the Kaplan-Meier estimate, and the solid curve represent model prediction. 210 The mean survival curve was within the 95% CI band of the estimated Kaplan-Meier curve.

211

Figure 10. Survival plot for cross-validation on the paediatric population. (Dotted curve

213 represents Kaplan–Meier estimate, and the solid curve represent model prediction) Cross validation on pediatric population: Age < 12</p>



214

215 External Validation

216 The external validation performed using DAISY data achieved a c-index 0.91 and 0.80 in years one and

two, respectively, even with a limited number of subjects, 40, in the external dataset (Table 24). The

c-index for subsequent years till six years was over 0.7. The VPC performed on the survival plot

showed good graphical fit (Figure 11). These results provide strong evidence for good predictive power for time frames over which a trial of reasonable duration would be conducted.

221 Figure 11. Survival plot for cross-validation on DAISY external validation dataset (dotted

222 curve represents Kaplan–Meier estimate and solid curve represents model prediction) External Validation using DAISY dataset



223

The survival modelling approach proposed by the applicant is overall consistent with previous
 recommendation and agreed upon in principle. The endpoint of interest (diagnosis of T1DM) is very
 well defined and usually non questionable from a clinical standpoint.

- However, several methodological issues were identified in the initial modelling implementation
- 228 approach as included in the initial proposal by the applicant, that were discussed during the DM, as 229 summarized below:

The applicant was invited to discuss the value of having a library of models included in the tool
 rather than a single model (as well as alternative approaches) to allow for flexibility in patient inclusion

criteria in the studies.

- In the briefing package, the applicant described the parametric AFT model. However, statistical
 notation and the description of the model was incorrect.
- The applicant suggested that covariates that were introduced in the model influence the
 baseline hazard (h0), which was only partially correct.
- The statistical notation in the original briefing document submitted by the applicant needed to
 be adjusted to better reflect this modelling approach. This inaccuracy has no influence on the
 presented simulations but is of importance when interpreting the estimated coefficients in the model
 (β).
- The applicant found that baseline age (bAGE_s) and SEX were highly associated with AA
 combinations (4.4.1.2. Analysis of Correlation and Association between Covariates). Hence, it was
 decided to not include bAGE_s and SEX in the subsequent multivariate analysis. This deserved
 additional justification.
- While it is acknowledged that the introduction of correlated covariates in a model can be
 problematic, especially when trying to predict in another dataset where this correlation between the
 covariates might be absent, it seems that the correlations between baseline Age and SEX and the AA
 combinations are similar for the TrialNet and TEDDY dataset. It also seems that adding SEX and
 baseline Age to the final AFT model would further reduce the AIC in a statistically significant manner.
- The consistency of covariate correlation across datasets was therefore crucial and it was
 requested that the applicant provides these data.
- Results of comparison of predictive performance of the proposed model with that of alternative
 models with other combinations of covariates were also requested, including a model with baseline Age
 and SEX in addition to the covariates identified by the applicant as final AFT model.
- Moreover, the prediction interval for the survival curves were missing and should be displayed
 in the figures, along with the R-code used to generate the VPCs that needed to be provided.
- As regards the statistical notation and the description of the model, the suggested modifications were implemented by the applicant. visual predictive check"-style figures and R code were provided as requested.
- 260 During the DM, in response to these issues, T1DC developed alternative models, including additional 261 variables: baseline age and sex. The original model improved when age and sex were included, as
- 262 indicated by the lower Akaike's Information Criteria (AIC) value. The time-dependent ROC curves and
- AUC values demonstrated good prediction performance (AUC > 0.75). Visual-predictive-check (VPC)style plots showed good graphical fit for internal and external validation of this selected model which
- 265 included age and sex.
- This was acknowledged by the qualification team (QT). It is considered important the applicant provides documented instructions to ensure the model is used correctly.
- 268 Alternative models were tested with different combinations of covariates including baseline age and
- sex in addition to the covariates previously included in the model. A table (table 20) was provided
- 270 showing the selected covariates for the alternative models. The predictive performance for these
- 271 models was compared using the AIC. The AIC value of alternative model 3 (alt_mod3) was significantly
- lower (with a reduction > 10) compared to all other alternative models and the original model. Hence,
- alternative model 3 (alt_mod3) was chosen as the selected model. Table 21 shows the parameter
- estimates for the selected model (alt_mod3).
- 275

276 Table 20. Value of AIC for the original model and other alternative models

| Model | Covariates | AIC |
|-----------------------------------|--|------|
| Original Model (orig_mod) | GAD65_IAA + GAD65_ZnT8 + IA-2_ZnT8 + IA- 2_IAA_ZnT8 + GAD65_IA-2_IAA_ZnT8+ Log_GLU120_s + HbA1c_s | 2982 |
| Alternative Model 1 (alt_mod1) | GAD65_IAA + GAD65_ZnT8 + IA-2_ZnT8 + IA- 2_IAA_ZnT8 + GAD65_IA-2_IAA_ZnT8+ Log_GLU120_s + HbA1c_s + SEX | 2972 |
| Alternative Model 2 (alt_mod2) | GAD65_IAA + GAD65_ZnT8 + IA-2_ZnT8 + IA- 2_IAA_ZnT8 + GAD65_IA-2_IAA_ZnT8+ Log_GLU120_s + HbA1c_s + bAGE_s | 2937 |
| Alternative Model 3 (alt_mod3) | GAD65_IAA + GAD65_ZnT8 + IA-2_ZnT8 + IA- 2_IAA_ZnT8 + GAD65_IA-2_IAA_ZnT8+ Log_GLU120_s + HbA1c s + bAGE s + SEX | 2921 |

277 Table 9. 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 |

278 Model performance for the selected model (alt_mod3) was assessed using time dependent Receiver

279 Operating Characteristic (ROC) curves and associated area under the curve (AUC) values (figure 12).

280 The internal validation for the selected model (alt_mod3) was performed using visual predictive check

281 (VPC)-style plots for a k-fold cross-validation and an internal validation with a paediatric population. An

external validation was performed with the DAISY dataset (Figures 9-11) and c-index values over 6
 years. The VPC-style plots overlaying observed data over model predictions showed good graphical fit.

284 The "survParamSim" package was used to generate the VPC-style plots.

Figure 12. Evaluation of model performance using time dependent receiver operation characteristic (ROC) analysis



Time dependent ROC analysis: model predictions on full analysis set

287



Figure 1. 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)



290

- 291 Figure 10. VPC-style plot for internal cross validation (CV) using pediatric population (red
- 292 shaded region shows the 95% prediction interval and the black shaded region shows the
- 293 95% CI for the observed data)



CV on pediatric population (age < 12) - alt_mod3

294

Qualification Opinion of Islet Autoantibodies (AAs) as Enrichment Biomarkers for Type 1 Diabetes (T1D) Prevention Clinical Trials EMA/580542/2021

295 Figure 11. VPC-style plot for external validation using the DAISY analysis dataset (red

296 shaded region shows the 95% prediction interval and the black shaded region shows the

297 95% CI for the observed data)





298

299 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. The AUC values for subsequent years for up to 5.5

- 301 years were greater than 0.75. These results provide evidence for good predictive power for time
- frames over which clinical trials of reasonable duration would be conducted. The c-index for the
 selected model (alt_mod3) for all five folds over six years was in most cases close to or higher than
- 304 0.8, suggesting good predictive performance.
- The alternative models developed by the applicant represent underlying evidence for the qualification of islet AAs as enrichment biomarkers and include clinically relevant glycaemic assessments (i.e., OGTT and HbA1c) as well as demographics (i.e., sex and baseline age) to allow for flexibility in patient inclusion criteria for T1D prevention studies. T1DC indicated that language will be drafted to guide
- sponsors to discuss with the regulatory agencies the use of this model to inform their drugdevelopment strategies.
- 311 Patient-level data from DAISY for the derived baseline showed similar distribution and correlation of
- covariates (including age, sex and AAs) compared to TEDDY and TN01 for the derived baseline. The
- 313 selected model showed adequate predictive performance across the three datasets for the selected
- covariates. The addition of age and sex improved model performance. T1DC indicated that they are
- open to continuing to test covariate correlation and updating the model as more data becomes
- 316 available.

317 Conclusion

- After the interactions with the SAWP, the applicant has provided a library of models, resulting in acceptable predictive performances for T1DM onset over a 6 years period.
- 319 It should be noted that additional covariates were also included in each of the proposed models beside
- 321 positivity to at least 2 Islet AAs. These additional predictors include HbA1c, blood glucose
- measurements from the 120-minute timepoints of an OGTT, baseline age and sex of patients. The
- 323 magnitude of the covariate effects for each of these predictors as well as their combination (OGTT,
- HbA1c, age and sex) was found to be higher than that of the Islet AAs alone. As a consequence, the
- 325 impact of the added-value of the positivity will for example be much less important for the patients
- 326 with already impaired OGTT (120-minute value between 7.8 and 11.1 mmol/L) and pre-diabetes
- 327 (fasting b-glucose 5.6 to 6.9 mmol/L).
- The use of the Islet AAs as a biomarker to optimize the design of clinical trials for the prevention of T1DM should therefore always be done also considering these additional patient characteristics.

330 **Question 4:**

331 EMA agree that the validation is adequate?

332 **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
- 334 for validation exactly once. This approach has been successfully used in prior qualification procedures
- 335 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
- 338 T1DC and used to perform patient-level external validation. This approach provided further evidence
- of robust model performance.

340 CHMP answer

- 341 VPC-style plots overlaying Kaplan-Meier curves over the selected model predictions showed good
- 342 graphical fit for folds 1, 2, 3 and 4 while fold 5 only performed well within the first year. For the
- internal cross validation using a paediatric population (age < 12), a c-index of 0.8 or higher was
- obtained until 3 years and a c-index of 0.75 or higher was obtained up to 6 years for the selected
- 345 model (alt_mod3) indicating good model performance. The visual predictive check (VPC) performed on
- the survival plot for cross-validation on the paediatric population (age < 12) showed reasonable
- 347 graphical fit. 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
- 349 (n=34). 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. The
 VPC performed on the survival plot showed good graphical fit (Figure 4).
- 352 It is agreed that these results provide strong enough evidence for good predictive power for time
- 353 frames over which a trial of reasonable duration would be conducted.
- 354 External validation was considered lacking in the qualification advice procedure. The applicant claims
- difficulty using published studies. The DAISY dataset was obtained for this purpose. In many ways it is
- 356 similar to the prior two datasets but, limited to one clinical centre over a long time period. The
- 357 numbers reaching the T1DM endpoint are low (N=19) compared to the other datasets. The clinical
- presentation (Table 2) differs significantly, with none of the patients developing DKA in DAISY. This
- 359 could be due to the small numbers but could also indicate other differences.

360 **Question 5:**

361 **Does EMA agree the presented results represent adequate supporting evidence for a** 362 **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 at least 2 islet AAs together with patient characteristics (sex, baseline age) and measures of glycaemic control (blood glucose measurements from the 120-minute timepoints of oral glucose tolerance test (OGTT)), and haemoglobin A1c (HbA1c) levels can help inform the definition of
- 369 entry criteria, enrichment strategies, and stratification approaches for T1D prevention clinical trials.

370 CHMP answer

- 371 The consortium has done much appreciated work in validating the model, in line with the feedback
- 372 from the regulators during the qualification advice and opinion procedures. A qualification is therefore
- 373 recommended for the positivity to at least 2 of the following islet AAs (IAA, GAD65, IA-2, and ZnT8),
- as measured using the analytical methods described in appendices A and B, and Addendum 1 as a
- biomarker of incidence of TD1M, when combined with the following additional baseline patients
- 376 characteristics of OGTT, HbA1c, age and sex.
- 377

378 3. Qualification opinion statement

- Positivity to at least 2 of the following islet AAs; IAA, GAD65, IA-2, and ZnT8 is qualified for use as enrichment biomarker, in combination with clinical parameters (sex, baseline age, blood glucose
- 201 enformment biomarker, in combination with chinical parameters (sex, baseline age, blood gidcose
- 381 measurements from the 120-minute timepoints of oral glucose tolerance test (OGTT), and hemoglobin
- A1c (HbA1c) levels) in T1D prevention trials targeting individuals at risk of developing T1D. A survival modelling approach was used to describe how the islet antibodies can be combined to the other patient
- 384 baseline characteristics for predicting timing to a T1D diagnosis.
- 385 The modeling exercise identified the relevance of additional clinical parameters (sex, baseline age,
- blood glucose measurements from the 120-minute timepoints of an oral glucose tolerance test (OGTT),and hemoglobin A1c (HbA1c) levels).
- 388 "At risk" was defined in this context as being a first degree relative (FDR) of a T1D patient, or having a
- 389 specific human leukocyte antigen (HLA) subtype of risk (HLA-DR3/3, DR4/4, DR3/4, DR3/X [X≠3],
- 390 DR4/X [X \neq 4]), excluding individuals with baseline fasting glucose \geq 126mg/dL (7.0 mmol/L) or
- 391 stimulated 2-hour glucose \geq 200 mg/dL (11.1 mmol/L).
- The present qualification opinion was requested for a new tool dedicated to enriching Type 1 Diabetes (T1D) prevention clinical trials. The proposed focus is on confirming the existence of a statistically
- 394 significant contribution of the positivity of of two or more islet autoantibodies (AAs) as predictors of
- 395 progressing towards a diagnosis of T1D, when combined with additional patient characteristics such as 396 OGTT, HbA1c, age and sex, as described in a validated survival model.
- 397 The applicant used an empirical/data driven modeling approach. In the absence of a mechanistic
- disease model, a clear and fully quantitative description of the contribution of the different factors
- including positivity to these AAs as predictors of progressing towards a diagnosis of T1D is therefore
- not possible. The models, as proposed, only allow confirming the existence of a statistically significant
 contribution of the different (combinations of) covariates and their relative relevance toward theT1D
- 401 contribution of the different (combinations of) covariates and their relative relevance toward theT1D402 diagnosis for patient at risk.
- 403 From a practical drug development standpoint, this proposed use is considered of added value because
- the intended application can help inform the definition of entry criteria, enrichment strategies, and
- stratification approaches in the field of T1D prevention. The clinical interest of identifying a good
- biomarker for Type 1 Diabetes (T1D) onset in an at-risk patient population is supported by the CHMP
- and the unmet need for better means to optimize drug development in the field is acknowledged.
- The model-based approach proposed by the applicant is considered an acceptable method to address the question of interest which is whether the combination of positivity to 2 or more of the 4 selected
- the question of interest which is whether the combination of positivity to 2 or more of the 4 selectedislet AAs can be considered acceptable predictors of a diagnosis of T1D, when combined to additional
- 411 and well-defined patient characteristics.
- 412 The analytical assays used to measure islet autoantibodies (AA) against glutamic acid decarboxylase
- 413 65 (GAD65), insulinoma antigen-2 (IA-2; also called ICA512), insulin (IAA), and zinc transporter 8
- 414 (ZnT8) in the three clinical studies contained in the modeling analysis are described in the 'key
- additional elements' section below. They are considered state of the art. It should be noted that the
- 416 results and the conclusions of the modeling analysis as assessed during this qualification procedure are
- 417 considered only applicable when the islet autoantibodies are measured using these methods or
- 418 methods proved to have at least equivalent analytical performances.
- 419 It should importantly be noted that this Qualification only refers to the value of the positivity of at least
- 420 two islet AAs in the risk assessment, when measured using the described analytical methods ('key
- 421 additional elements' section below), or methods with comparable accuracy, sensitivity and specificity.
- 422 The data used for the model development and external validations to support the qualification of islet
- 423 AAs as enrichment biomarkers originated from three independent datasets: The Environmental
- 424 Determinants of Diabetes in the Young (TEDDY), the TrialNet Pathway to Prevention Study (TN01) and
- the Diabetes Autoimmunity Study in the Young (DAISY) the TN01, TEDDY, and DAISY registry studies.
- 426 Details are provided in the answer to Question 2 by the applicant.

- 427 The data sources are judged largely relevant, consistent with the recommendation during the QA
- 428 procedure. From a modeling perspective, this approach is endorsed, and the 3 data sources seem
- 429 adequate. Potential covariate distribution and correlation were presented and discussed as requested430 during the qualification procedure.
- 431 The baseline data intended for modeling are relatively well defined, as well as the binary endpoint (T1D
- diagnosis). Longitudinal assessments of islet AA positivity, OGTTs, C-peptide measurements, and
- 433 HbA1c measurements are considered out of scope for the proposed analysis, and only baseline
- 434 information were used for the modeling analysis.
- The precise definition of baseline used for the analysis set is the first record, (i.e., time point) for each individual in which the following criteria is satisfied:
- 437 Presence of any two or more of the 4 islet AAs
- Complete, (i.e., non-missing) information for OGTT (0 and 120-minute time points), HbA1C
 measurements, age and sex.
- The applicant developed a survival model to describe the time course of incidence of T1DM in patients
- included in the 2 datasets used for model building (TEDDY and TN01), given their baseline
- 442 characteristics. The third dataset (DAISY) was used for model validation.
- 443 The details and different steps of modeling methodology, model development, internal and external
- validation are described as initially provided by the applicant in answer to Questions 3 and 4. Following
- the DM with SAWP, several components of the proposed modelling plan were updated according to
- SAWP feedback. The updated modelling analysis plan was executed, submitted to SAWP, and discussedat a subsequent DM.
- Briefly, the applicant has provided a library of models, including a preferred selected model, resulting
- in acceptable predictive performances for T1DM onset over a 6-year period. It should be noted that
- additional covariates were also included in each of the proposed models beside the positivity to at least
 2 islet AAs. These additional predictors are HbA1c, blood glucose measurements from the 120-minute
- 451 timepoints of OGTT, baseline age and sex. The magnitude of the covariate effects for each of these
- 453 predictors as well as their combination (OGTT, HbA1c, age and sex) was found to be higher than that
- 454 of the IAAs. As a consequence, the impact of the added-value of the positivity will for example be
- 455 much less important for the patients with already impaired OGTT (120-minute value between 7.8 and
- 456 11.1 mmol/L) and pre-diabetes (fasting b-glucose 5.6 to 6.9 mmol/L).
- The models that provided the evidence for this qualification opinion are available in the 'key additional elements' section below as implemented in R software (The R code used to implement the model with the best predictive performances is provided below).
- 460 In conclusion, the use of the islet AAs as biomarkers to optimize the design of clinical trials for the
- 461 prevention of T1DM should therefore always be done also considering these additional patient
- 462 characteristics, as described in the models.
- 463 **4. Key additional elements**

464 **4.1.** Islet autoantibody analytical assays

465 General background on Islet autoantibody assays

- 466 Multiple assays were used to measure islet autoantibodies (AA) against glutamic acid decarboxylase 65 467 (GAD65), insulinoma antigen-2 (IA-2; also called ICA512), and insulin (IAA) in the two clinical studies 468 contained in our modeling analysis: TrialNet Pathway to Prevention (TN01TN01, formally called TrialNet 469 natural history study), and The Environmental Determinants of Diabetes in the Young (TEDDY). Data 470 were collected over the period of 2004-2018 for TN01 and 2004-2016 for TEDDY. These dates were 471 generated by reference laboratories in Denver, CO (Barbara Davis Center, University of Colorado; Dr. 472 Liping Yu is Principal Investigator) for both TN01 and TEDDY studies and in Bristol, UK (University of 473 Bristol; Mr. Alistair Williams is Principal Investigator) for only TEDDY study. Both laboratories have 474 extensive experience in immunoassay development and validation with a strong record of publication 475 in peer-reviewed journals. The laboratory in Denver was CLIA certified in 2000. Islet AAs were
- 476 measured in serum using standardized radio-binding assays (RBAs) whose methodological details have

- 477 been published. A sample is determined as "positive" or "negative" for a particular islet AA according to
- 478 pre-specified thresholds determined with reference samples (i.e., sera from patients with recently
- diagnosed with T1D diabetes as positives, and sera from normal patients as negatives). In addition,

483

values.

- robust procedures were used by both laboratories to ensure accuracy of positive calls and the
 consistency of responses over time. Emphasis for this EMA submission document is placed on a binary
 determination of seropositivity or seronegativity of islet AAs, rather than on quantitation of continuous
- Prior to 2010, data were generated using "local" assays developed and published by the Denver and 484 485 Bristol laboratories. However, starting in 2010, both laboratories implemented NIDDK sponsored 486 "harmonized" assays for autoantibodies to GAD65 and IA-2 (but not for ZnT8 or IAA autoantibody 487 assays) which were developed under the direction of the Islet Autoantibody Harmonization Committee, 488 which included the use of common reference standards (for generating standard curves and common 489 units of autoantibody levels in serum) from the US National Institute of Diabetes and Digestive and 490 Kidney Diseases (NIDDK). This project was also supported by the Islet Autoantibody Standardization 491 Program (IASP), formerly known as the Diabetes Autoantibody Standardization Program (DASP), which 492 is an international effort to improve and harmonize measurement of islet AAs associated with T1D 493 through proficiency testing, and by providing advice, training, and information. The Centers for Disease 494 Control and Prevention (CDC) have participated in this National Institutes of Health (NIH) sponsored 495 standardization effort. Every 18 months IASP carries out a voluntary or opt-in assessment program for 496 labs around the world that perform islet AA assays. In this assessment, IASP provides between 50-150 497 blinded seropositive and seronegative sera samples sets from T1D patients and control subjects as well 498 as reference standard reagents to participating laboratories, and the results released to laboratories to 499 continually compare and improve assay performance. Data from the DASP/IASP assessments for the 500 Bristol and Denver labs are described later in this document under the discussion of concordance. 501 The qualitative, binary determination of seropositivity or seronegativity for each islet autoantibody is a 502 key feature in the modeling plan outlined in Section 4.3.1 of the Briefing Document. Calling a particular 503 sample positive for a given autoantibody is defined as when the measured value exceeds a cutoff that 504 was set at an antibody prevalence in reference populations of healthy individuals and those with T1D. 505 Ideally, the reference populations should have similar characteristics to the at-risk population and be 506 large enough to achieve tight confidence intervals. For the determination of positivity cutoffs, positive 507 controls are serum samples from patients newly diagnosed (within two weeks) with T1D, and negative 508 controls are serum samples from healthy individuals. The cutoff is commonly set at the 99th percentile 509 of the reference population, i.e. a level exceeded by only 1% of these healthy individuals. For the 510 GAD65 and IA-2 harmonized assays (i.e., from 2010 onwards) from Denver and Bristol, NIDDK 511 standards were provided to establish a six-point standard curve for the calculation of standardized 512 Digestive and Kidney (DK) units that were then compared to pre-specified cutoffs for determination of 513 seropositivity or negativity. These NIDDK standards were run in each assay and were provided as part 514 of the harmonization program. For all IAA assays run in Denver, and for GAD65 and IA-2 assays prior 515 to 2010 (termed "local" assays), positive control sera from newly diagnosed T1D patients and negative 516 control sera from healthy subjects were used by the Denver lab to generate an index that enabled the 517 determination of seropositivity or negativity. The index is a ratio of the signal in the test serum to the 518 signal in a positive control; if that ratio exceeds the pre-specified cutoff, then the sample is called 519 seropositive. In the GAD65 and IA-2 assays run before 2010 in Bristol, locally prepared standards were 520 used to generate standard curves for the calculation of World Health Organization (WHO) units that 521 were then compared to pre-specified cutoffs for determination of seropositivity or negativity. In 522 addition, a detailed discussion of how seropositivity was confirmed can be found in Section 4.3 of the 523 Briefing Package. 524
- The assays for GAD65 and IA-2 AAs that generated data for this submission are not quantitative and are only being used in this submission to determine the presence or absence of an individual AA. Some of the features of these islet AA assays that prevent them from being used quantitively are:

The amount of radio-labeled antigen generated in the *in vitro* transcription/translation reaction
 is not quantitated.

• The radio-labeled antigen does not fully saturate binding sites of the serum AAs.

• There is no step to compete off non-specific binding using excess cold antigen.

531 For these reasons, the absolute lower limit of quantitation (LLOQ) and upper limit of quantitation 532 (ULOQ) are not determined for these assays. In addition to the points stated above, because the 533 autoantibodies being detected are a composition of polyclonal antibodies that differ in affinity and 534 concentration, parallelism studies and linearity assessments have not been performed. Although these 535 factors prevent the use of the continuous measure from these islet AA assays, robust positive and 536 negative controls enable the binary adjudication of seropositivity or negativity. Similarly, the Islet AA 537 assay as performed by the Denver and Bristol labs is not quantitative, despite the ability to quantify 538 the antigen and the inclusion of unlabeled insulin to reduce non-specific binding. Although quantitative 539 IAA assays could be developed, those used in this submission were not run in a quantitative format 540 and only the binary output is being utilized.

541 Although FDA 501k-cleared assays are available to measure some of the AAs, samples for the two 542 studies were analyzed in two different laboratories using either local or harmonized radiobinding 543 assays (RBAs) that were published by the participating laboratories as summarized in Table 1 of this 544 document. This assay format is commonly used for measurement of AAs because it is high throughput, 545 relatively inexpensive, uses small serum volumes, is easily adapted for detection of different AAs (by 546 changing the radiolabeled antigen) and performed better than other immunoassays such as ELISA 547 because of the RBAs solution phase format that facilitates antigen-antibody binding. Should sponsors 548 want to measure islet AAs in future clinical studies, they may choose to use different assays, including 549 those that do not require radiolabels. To verify that these future assays are indeed fit for purpose, a 550 proficiency test consisting of a panel of samples comprising different levels of islet AAs should be 551 performed. This proficiency test would evaluate the same panel of 7 samples in both the RBAs 552 described here and these future alternative assays. This proposed proficiency test is not discussed any 553 further as it is not the focus of this submission. Users of any proposed future islet AAs assay will be 554 required to provide detailed information on precision and relative accuracy.

As with the assays for GAD65, IA-2, and insulin AAs, the ZnT8 AA assay that generated data for this submission is not quantitative and is only being used in this submission to determine the presence or absence of an individual AA. Some of the features of these islet AA assays that prevent them from being used quantitively are:

- The amount of radio-labeled antigen generated in the *in vitro* transcription/translation reaction
 is not quantitated.
- The radio-labeled antigen does not fully saturate binding sites of the serum AAs.
- There is no step to compete off non-specific binding using excess cold antigen.

563 For these reasons, the absolute lower limit of quantitation (LLOQ) and upper limit of quantitation 564 (ULOQ) are not determined for these assays. In addition to the points stated above, because the AA 565 being detected are a composition of polyclonal antibodies that differ in affinity and concentration, 566 parallelism studies and linearity assessments have not been performed. Although these factors prevent 567 the use of the continuous measure from these islet AA assays, robust positive and negative controls 568 enable the binary adjudication of seropositivity or negativity. 569 Samples were analyzed using a local radiobinding assay (RBA) assay format that is commonly used for 570 measurement of AAs because it is high throughput, relatively inexpensive, uses small serum volumes, 571 and is easily adapted for detection of different AAs (by changing the radiolabeled antigen). In addition, 572 the assay performed better than other immunoassays such as ELISA because of the RBA's solution-573 phase format that facilitates antigen-antibody binding. Should sponsors want to measure islet AAs in

- 574 future clinical studies, they may choose to use different assays, including those that do not require
- 575 radiolabels. To verify that these future assays are indeed fit for purpose, a proficiency test consisting of
- a panel of samples comprising different levels of islet AAs should be performed. This proficiency test

577 would evaluate the same panel of samples in both the RBAs described here and these future

alternative assays. This proposed proficiency test is not discussed any further as it is not the focus of

579 this submission. Users of any proposed future islet AA assays will be required to provide detailed

580 information on precision and relative accuracy.

581 **Table 1. Autoantibody assay summary**

| | Trial/study | Site | | Assay |
|--------------|-------------|------------|-------------------|-------------------|
| Autoantibody | name | Measured* | RBA Assay Type** | Documentation |
| | | | | |
| GAD65 | TN01 | UC | Local, Harmonized | 2019 Briefing Pkg |
| GAD65 | TEDDY | UC/Bristol | Local, Harmonized | 2019 Briefing Pkg |
| GAD65 | DAISY | UC | Local, Harmonized | 2019 Briefing Pkg |
| | | | | |
| IA-2 | TN01 | UC | Local, Harmonized | 2019 Briefing Pkg |
| IA-2 | TEDDY | UC/Bristol | Local, Harmonized | 2019 Briefing Pkg |
| IA-2 | DAISY | UC | Local, Harmonized | 2019 Briefing Pkg |
| | | | | |
| IAA | TN01 | UC | Local | 2019 Briefing Pkg |
| IAA | TEDDY | UC/Bristol | Local | 2019 Briefing Pkg |
| IAA | DAISY | UC | Local | 2019 Briefing Pkg |
| | | | | |
| ZnT8 | TN01 | UC | Local | 2020 Update |
| ZnT8 | TEDDY | UC | Local | 2020 Update |
| ZnT8 | DAISY | UC | Local | 2020 Update |
| | | | | |

582 * UC = UC Core Facility

** For GAD65 and IA-2, local assays were used for samples analyzed before 2010 and harmonized
 assays were used for samples analyzed starting in 2010.

In summary, the assays used to generate the islet AA data were performed in central laboratories that have been participating in multi-center diabetes studies and international islet AA assay harmonization workshops for more than 20 years and the methodologies for all assays have been published in peerreviewed journals. Importantly, robust procedures, including the use of QC controls that have shown strong concordance between labs and minimal variability over time, were used by both laboratories to ensure accuracy of positive calls or seropositivity or seronegativity and the consistency of responses over time.

592 Summary of GAD65 and IA-2 AA assays

593 **Overview**

594 Radiobinding assays are used to qualitatively determine the presence or absence, as seropositivity or 595 seronegativity, of the AAs to GAD65 and IA-2 (the local Denver IA-2 was originally called islet cell 596 antigen 512 [ICA512]) in serum samples from patients. For most of these RBAs, one autoantibody is 597 assessed per well (i.e., using one radiolabeled antigen), except for the local GAD65 and IA-2 assays in 598 Denver that are multiplexed using different radiolabels for each antigen. In this assay format, in vitro 599 transcription and translation (IVTT) is used to generate a specific radiolabeled human antigen (either 600 GAD65 or IA-2) using a radiolabeled amino acid in rabbit reticulocyte lysates. Once prepared, the 601 radio-labeled antigen is incubated with patient serum overnight. A non-specific immunoglobulin 602 precipitation is then carried out with Protein A-Sepharose beads to isolate radiolabeled antigen-islet 603 autoantibody complex to enable removal of unincorporated radiolabeled amino acids. The washed, 604 isolated beads are then assessed via scintillation counting to evaluate the levels of radiolabeled antigen 605 that have been isolated. These levels are then compared to positive controls for the determination of 606 seropositivity.

Figure 1: Schematic of Radiobinding Assay format used for GAD65 and IA-2 AA assessments



607

608 The harmonized GAD65 and IA-2 autoantibody assays both use calibrators/standards developed by the 609 NIDDK that were prepared from a set of positive and negative calibrators prepared from pooled sera 610 (Bonifacio et al. 2010). For the positive calibrator, 25–50 ml serum was collected from each of 21 611 patients with T1D aged 18 to 30 years with a median time since diagnosis of 1.1 year (range, 0.2 to 612 2.2 year). The presence of moderate/high levels of antibodies to GAD and IA-2 in individual sera was 613 confirmed in the Denver, Bristol, and Munich (a laboratory from Helmholtz Zentrum München 614 participated in harmonization efforts but did not assay samples for TEDDY and TN01) laboratories 615 before the sera were pooled. For the negative serum diluents and calibrator, 12 frozen serum 616 donations (median volume for each sample, 228 ml) were obtained from the blood bank of the Städt 617 Klinikum München GmbH, Munich and pooled.

- 618 While there are general similarities in how the different RBAs are performed for all autoantibodies, 619 there are also differences when comparing the local and harmonized assays from the same site, as well 620 as the assays from Denver and Bristol. For example, each site prepares its own local QC standards that 621 are used to assess performance over time and to ensure the assay is functioning properly. Other
- differences specific for GAD65 and IA-2 are highlighted in the text and tables below.

623 GAD65 and IA-2 Assay Characteristics

624 GAD65: In comparing the local assays from Denver and Bristol, there are several differences. The 625 local Denver assay measures GAD65 in a multiplexed format with IA-2 (called ICA512 in the SOP) in 626 which GAD65 is labeled with 3H-leucine and IA-2 is labeled with 35S-methionine in separate IVTT 627 reactions and then the two labeled antigens are mixed with the serum in the assay. Also, the Denver 628 assay uses an index for determining seropositivity whereas the Bristol assay uses WHO units, which 629 were established at the first DASP workshop in 2000 (Mire-Sluis et al. 2000). All versions of the GAD65 630 assay used expression plasmids encoding the full-length protein. In comparing the harmonized assays, 631 the methods are highly similar, but as mentioned, different local QC controls are used. Table 2 632 compares the local and harmonized Denver and Bristol GAD65 assays. In addition, only the Bristol lab 633 uses a confirmatory threshold (20 DK units, which is set below the positivity threshold of 33 DK units 634 to avoid introducing a negative bias); samples that exceed the threshold are repeated in a separate 635 assay and reported as the mean of the two results. Finally, the positivity cutoff for the harmonized 636 assay run in Denver is 20 DK units/ml, whereas it is 33 DK units/ml for the Bristol assay.

Table 2. Comparison of local and harmonized GAD65 assays from Denver and Bristol

| | GAD65 | | | |
|------------------------|------------------------|----------------------------|----------------------------|----------------------------|
| Local or Harmonized | Local | Local | Harmonized | Harmonized |
| Site | Denver | Bristol | Denver | Bristol |
| Antigen source | IVTT | IVTT | IVTT | IVTT |
| Plasmid clone | pEx9-GAD65 | pGEM3-GAD65 | pTH-GAD65 | pTH-GAD65 |
| Source of plasmid | A. Lernmark | E. Bonifacio | A. Lernmark | A. Lernmark |
| Plasmid Reference | <u>Grubin 1994</u> | Bonifacio 1995 | <u>Hansson 2010</u> | Hansson 2010 |
| Amino acids expressed | Full length | Full length | Full length | Full length |
| Local OC controls | High pos, low | High pos, med | High pos, low | High pos, med pos, |
| Local QC controls | pos, neg | pos, low pos, neg | pos, neg | low pos, neg |
| Colibuatou (Standauda | Same as QC | | | |
| Calibrator / Standards | controls | Locally prepared | 7 from NIDDK | 7 from NIDDK |
| Radiolabel | ³ H-Leucine | ³⁵ S-Methionine | ³⁵ S-Methionine | ³⁵ S-Methionine |
| Sepharose beads | Protein A | Protein A | Protein A | Protein A |
| Multiplexed | Yes (with IA-2) | No | No | No |
| Assay Units | Index | WHO units/ml | DK units/ml | DK units/ml |

637

638 IA-2: In comparing the local assays from Denver and Bristol, there are several differences. The local 639 Denver assay measures IA-2 (called ICA512 in SOP) in a multiplexed format in which the IA-2 is 640 labeled with 35S-methionine and GAD65 is labeled with 3H-leucine in separate IVTT reactions and then 641 the two labeled antigens are mixed with the serum in the assay. Also, as with GAD65, the local Denver 642 assay uses an index for determining seropositivity whereas the Bristol assay uses WHO units, which 643 were established at the first DASP workshop in 2000 (Mire-Sluis et al. 2000). Finally, the antigen 644 expressed in the local Denver IA-2 assay (pCRII-ICA512bdc, amino acids 256-979) is different from 645 local Bristol assay (pSP64 IA-2ic, 605 to 979) and the antigen in the harmonized assay (pSP64-PolyA-646 IA-2ic, amino acids 606 to 979). Table 3 compares the local and harmonized Denver and Bristol IA-2 647 assays.

648 To demonstrate alignment between the local Denver IA-2 assay, which utilized a long-form construct 649 (256-979), and the Bristol local and harmonized assays, which used a shorter construct (606 to 979), 650 a comparison carried out between both labs was performed using 2,172 TN01 samples. These 2,172 651 samples included: 1,089 samples positive for any AA with the "local" TrialNet assays (GAD65, IAA and 652 IA-2) and 1,074 randomly selected antibody negative samples. The local Denver IA-2 and harmonized 653 assays from Denver were 95% concordant for positives or negatives with r2= 0.72 for IA-2 AAs. In 654 comparing the harmonized assays from Bristol and Denver, the methods are highly similar, but as 655 mentioned, different local QC controls are used. In addition, only the Bristol lab uses a confirmatory 656 threshold (1.4 DK units, which is set below the positivity threshold of 5 DK units to avoid introducing a 657 negative bias); samples that exceed the threshold are repeated in a separate assay and reported as 658 the mean of the two results.

Table 3. Comparison of local and harmonized IA-2 assays from Bristol and Denver.

| | IA-2 | | | |
|------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| Local or Harmonized | Local | Local | Harmonized | Harmonized |
| Site | Denver | Bristol | Denver | Bristol |
| Antigen source | IVTT | IVTT | IVTT | IVTT |
| Discription | | | pSP64-PolyA-IA- | pSP64-PolyA-IA- |
| Plasmid clone | pCRII-ICA512bdc | pSP64 IA-2ic | 2ic | 2ic |
| Source of placmid | Barbara Davis | | | |
| Source of plasmid | Center | M. Christie | Ezio Bonifacio | V. Lampasona |
| Plasmid Reference | <u>Gianani 1995</u> | Hatfield 1997 | Bonifacio 2010 | Bonifacio 2010 |
| Amino acids expressed | 256-979 | 605-979 | 606-979 | 606-979 |
| Less OC controls | High pos, low pos, | High pos, med | High pos, low | High pos, med |
| Local QC controls | neg | pos, low pos, neg | pos, neg | pos, low pos, neg |
| Calibrator /Standardo | Same as QC | | | |
| Calibrator / Standards | controls | Locally prepared | 7 from NIDDK | 7 from NIDDK |
| Radiolabel | ³⁵ S-Methionine | ³⁵ S-Methionine | ³⁵ S-Methionine | ³⁵ S-Methionine |
| Sepharose beads | Protein A | Protein A | Protein A | Protein A |
| Multiplexed | Yes (with GAD65) | No | No | No |
| Assay Units | Index | WHO units/ml | DK units/ml | DK units/ml |

660

671

661 3 Summary of the Insulin AA Assay

662 **Overview**

663 For detection of insulin AAs, 125I-insulin is used as the antigen rather than in vitro transcription and 664 translation (Figure 2). There is also an additional step that includes competition with unlabeled ("cold") 665 insulin (to reduce non-specific binding), and immunoglobulin-binding Sepharose beads are used to 666 isolate the radiolabeled antigen-islet AA complex to enable removal of unincorporated radiolabeled 667 amino acids. In parallel, samples are incubated with either 125I-insulin alone, or with a combination of 668 125I-insulin and cold insulin, and the results are calculated based on the difference in radioactivity 669 between the two for each sample. In all assays, an islet AA is called positive, if the measurement in 670 the assay exceeds a predefined positivity threshold/cutoff.

Figure 2: Schematic of Insulin AA Radiobinding Assay Format



672 Insulin AA Assay Characteristics

- In Bristol, the assay is run in two stages: first, a screening assay (IAA) in which samples are tested for
- 674 insulin binding using 125I-insulin alone (hot label) is run; if above the screening threshold then a
- 675 competition assay (CIAA) is run in which specificity of insulin binding is confirmed by displacement of

676 binding to 125I-insulin label by addition of excess recombinant insulin (cold label). The CIAA results 677 are calculated using the mean delta cpm (cpm with hot label – cpm with cold label) for each sample. 678 The Denver assay is run similarly, except that, starting in 2015, if the signal of the duplicate with cold 679 insulin is \geq 20 cpm, then one of two next steps is taken: If the count is less than 1,000 cpm and is 680 confirmed by a 2nd run, then the sample will be considered not reportable (due to non-specific 681 binding). However, if the counts are greater than or equal to 1,000 cpm and confirmed by 2nd run, 682 then the assay will be re-run with 10x more cold insulin. The reason for this two-step process is that in 683 rare instances, some samples gave false positive signals in the original IAA assay. It was determined 684 that these false positives were due to the presence of 125I-labeled peptides other than insulin, which 685 were related to insulin and present as a contaminant of the purchased 125I-insulin reagent. This was 686 demonstrated because the false positive signal could not be competed with cold insulin and was hence 687 likely due to the presence of non-insulin 125I-labeled peptides being present in the purchased 125I-688 insulin reagent.

689 The Denver assays uses a combination of Protein A-Sepharose and Protein G-Sepharose, whereas the 690 Bristol lab only uses Protein A-Sepharose. The determination of positivity in the Denver assay is based 691 on an index using local QC controls, whereas a standard curve of arbitrary units is used for the Bristol 69

| 92 | assay. | Table 5 | compares | the local | Denver | and Bristo | IAA assays. | |
|----|--------|---------|----------|-----------|--------|------------|-------------|--|
| | | | | | | | | |

| Local or Harmonized | Local | Local |
|-----------------------|---------------------------|--------------------------|
| Site | Denver | Bristol |
| Antigen source | Amersham | Amersham |
| Plasmid clone | NA | NA |
| Source of plasmid | Perkin Elmer | Perkin Elmer |
| Plasmid Reference | NA | NA |
| Amino acids expressed | Full length | Full length |
| Local OC controls | High pos, low pos, ultra- | High pos, med pos, |
| Local QC controls | low pos, neg | low pos, neg |
| Calibrator/Standards | Same as QC controls | Locally prepared |
| Radiolabel | 1251-Insulin | ¹²⁵ I-Insulin |
| Sepharose beads | Protein A and Protein G | Protein A |
| Multiplexed | No | No |
| Assay Units | Index | Arbitrary units |

Table 5. Comparison of local IAA assays from Denver and Bristol

693

694 **ZnT8** Assays

695 Data from the ZnT8 assay were generated by the Autoantibody/HLA Core Facility at the University of

696 Colorado (UC), Aurora, CO, USA; (referred to as the UC Core Facility throughout the rest of this

697 document and referred to as the "Denver lab"). Islet AAs were measured in serum using standardized

698 radio-binding assays (RBAs) whose methodological details have been published [1, 2]. A sample is

699 determined as "positive" or "negative" for a particular islet AA according to pre-specified thresholds

700 determined with reference samples (i.e., sera from patients with recently diagnosed with T1D diabetes

- 701 as positives, and sera from normal patients as negatives).
- 702 1. Lampasona V, Schlosser M, Mueller PW, et al (2011) Diabetes Antibody Standardization Program: 703 First Proficiency Evaluation of Assays for Autoantibodies to Zinc Transporter 8. Clinical Chemistry 704 57(12):1693-1702. https://doi.org/10.1373/clinchem.2011.170662
- 705 2. Yu L, Herold K, Krause-Steinrauf H, et al (2011) Rituximab selectively suppresses specific islet 706 antibodies. Diabetes 60(10):2560-2565. https://doi.org/10.2337/db11-0674
- 707

708 R code for the final model (i.e. with the best predictive performances) 709 710 R markdown file number: "4" 711 title: "Model validation - Islet AA for EMA qualification" 712 author: "T1DC modeling team at C-Path" 713 last updated: 12 May 2020 714 715 This R markdown file contains code for model validation including K-fold and external validation with 716 DAISY dataset. The result from running a code block can be viewed under the code block. Additionally, 717 the figures and tables generated from these code blocks will be saved in "figures" and "tables" folder 718 under "deliv" folder. The associated file names for the figures and tables describe the data being 719 visualized or tabulated. ```{r Check if relevant libraries are installed on local machine, install otherwise} 720 #Function to check whether a package is installed 721 722 is.installed <- function(mypkg) {</pre> is.element(mypkg, installed.packages()[, 1]) 723 724 } 725 #A tool for fast aggregation of large data 726 if (is.installed("data.table") == FALSE) { 727 install.packages("data.table", dependencies = TRUE) 728 } 729 #A library for computing survival analyses 730 if (is.installed("survival") == FALSE) { 731 install.packages("survival", dependencies = TRUE) 732 } 733 #A library for visualizing survival analysis results 734 if (is.installed("survminer") == FALSE) { 735 install.packages("survminer", dependencies = TRUE) 736 } 737 #A library of r packages to perform data science tasks 738 if (is.installed("tidyverse") == FALSE) { 739 install.packages("tidyverse" , dependencies = TRUE) 740 } 741 #A package to generate correlation plots 742 if (is.installed("corrplot") == FALSE) { 743 install.packages("corrplot", dependencies = TRUE) 744 } 745 #A package to perform survival analysis 746 if (is.installed("flexsurv") == FALSE) { 747 install.packages("flexsurv", dependencies = TRUE) 748 } 749 #A package to compute time-dependent ROC curve from censored survival data if (is.installed("survivalROC") == FALSE) { 750 751 install.packages("survivalROC", dependencies = TRUE) 752 } 753 #A toolbox for assessing and comparing performance of risk predictions 754 if (is.installed("riskRegression") == FALSE) { 755 install.packages("riskRegression", dependencies = TRUE) 756 } 757 #A package for estimation of prediction accuracy for time-to-event data 758 if (is.installed("survAUC") == FALSE) {

| 759 | install.packages("survAUC" , dependencies = TRUE) |
|-----|---|
| 760 | } |
| 761 | • • • |
| 762 | ```{r load libraries} |
| 763 | library(data.table) #A tool for fast aggregation of large data |
| 764 | library(survival) #A library for computing survival analyses |
| 765 | library(survminer) #A library for visualizing survival analysis results |
| 766 | library(tidyverse) #A library for r packages for perform data science tasks |
| 767 | library(corrplot)#A package to generate correlation plots |
| 768 | library(flexsurv)#A package to perform survival analysis |
| 769 | library(survivalROC)#A package to compute time-dependent ROC curve from censored survival data |
| 770 | library(riskRegression) #A toolbox for assessing and comparing performance of risk predictions |
| 771 | library(survAUC) #A package for estimation of prediction accuracy for time-to-event data |
| 772 | #library(rms) |
| 773 | |
| 774 | ```{r Clear environment} |
| 775 | rm(list=ls()) |
| 776 | |
| 777 | ```{r Load modeling analysis datasets generated from R markdown file 1 from the "data" folder} |
| 778 | #Model analysis dataset from TN01 and TEDDY |
| 779 | data <- readRDS("/data/final_EMA_islet_AA_datamart.rds") |
| 780 | #External validation dataset from DAISY |
| 781 | data_daisy <- readRDS("/data/final_EMA_daisy_datamart.rds") |
| 782 | *** |
| 783 | ```{r Recode subject IDs to be consecutive integers} |
| 784 | data\$IDp <- data\$IDp_new |
| 785 | *** |
| 786 | ```{r K-fold cross-validation analysis as discussed in section 4.3.7.2 - generating random k-folds} |
| 787 | #Set a seed value for random split |
| 788 | set.seed(1) |
| 789 | #set number of folds to 5 |
| 790 | n <- 5 |
| 791 | #Generate 5 random data splits |
| 792 | cv <- getSplitMethod(paste0("cv",n), B=1, N=2022) |
| 793 | folds <- cv[[3]] |
| 794 | folds <- as.factor(folds) |
| 795 | splits <- split(data, folds) |
| 796 | *** |
| 797 | ```{r K-fold cross-validation analysis as discussed in section 4.3.7.2} |
| 798 | #Set a seed value |
| 799 | set.seed(1) |
| 800 | #Assign maximum year for c-index calculation |
| 801 | yrs_for_cindex <- 6 |
| 802 | #Create a matrix to store c-index values |
| 803 | cindex_k_fold <- matrix(NA,nrow = n, ncol = yrs_for_cindex) |
| 804 | #Apply for loop to rotate folds for cross-validation |
| 805 | for(i in 1:n){ |
| 806 | |
| 807 | train <- data.frame() |
| 808 | train_inds <- c(1:n) |
| 809 | train_inds <- train_inds[-i] |

```
810
        test ind <- i
811
        for(j in 1: (n-1)) {train <- rbind(train,splits[[train_inds[j]]])}</pre>
812
        test <- splits[[test_ind]]</pre>
813
814
815
          #Fit model using 'flexsurvreg" function with final multivariate AFT model described in section
816
       4.4.2.4
817
          surv_obj_train <- Surv(train$T_event, train$status)</pre>
818
          fit_train <- do.call(flexsurvreg, list(formula = surv_obj_train ~ GAD65_IAA + GAD65_ZNT8 +
819
       IA2A_ZNT8 + IA2A_IAA_ZNT8 + GAD65_IA2A_IAA_ZNT8 + HbA1c_s +
                                                                                   log_GLU120_s, data =
820
       train, dist = "Weibull"))
821
822
          #Use "survreg" to compute c-index
823
          fit_train_concordance<- survreg(Surv(T_event, status) ~ GAD65_IAA + GAD65_ZNT8 +
824
       IA2A_ZNT8 + IA2A_IAA_ZNT8 + GAD65_IA2A_IAA_ZNT8 + HbA1c_s + log_GLU120_s , data =
825
       train .dist = "weibull" )
826
          #Check model fit with test fold
827
          fit_test <- survfit(Surv(T_event, status) ~ 1, data = test)</pre>
828
          surv1 <- summary(fit train, newdata = test, type = "survival", B=1, tidy = TRUE)
829
          varnames <- c("time", "surv", "lower", "upper")</pre>
830
          fit_test_data <- cbind(fit_test$time, fit_test$surv, fit_test$lower, fit_test$upper)
831
          fit test data <- as.data.frame(fit test data)
832
          names(fit_test_data) <- varnames</pre>
833
834
          surv avg <- surv1 %>%
835
           group by(time) %>%
836
           summarise(mean_est = mean(est, na.rm=TRUE),
837
                  mean lcl = mean(lcl, na.rm=TRUE),
838
                  mean ucl = mean(ucl, na.rm=TRUE),
839
                  )
840
          #Generate plot to check goodness-of-fit
841
          p < -qqplot() +
842
          qqtitle(paste("Cross validation on Fold ",i, sep = "")) +
843
          geom line(data = surv avg, aes(x = time, y = mean est)) +
844
          geom_step(data = fit_test_data, aes(x = time, y = surv), linetype = 3, size = 1) +
845
          geom_ribbon(data = fit_test_data, aes(x = time, ymin = lower, ymax = upper), linetype = 0,
846
       alpha = .2, show.legend = FALSE) +
847
          geom_ribbon(data = surv_avg, aes(x = time, ymin = mean_lcl, ymax = mean_ucl), linetype = 0,
848
       alpha = .2, show.legend = FALSE) +
849
          xlab("Time from Derived BL (years)") +
850
          ylab("1 - Probability of T1D Diagnosis")
851
852
          #View goodness-of-fit plot
853
          р
854
855
          #Export cross-validation plots
856
          ggsave(paste("../deliv/figures/",i," fold_validation",".png", sep = ""), p, width = 16, height = 9,
857
       units = "cm")
858
         #Compute c-index for model prediction on kth fold
859
          for(q in 1:yrs_for_cindex){
```

```
860
            c index tmp <- concordance(object = fit train concordance, newdata = test, ymin = 0,ymax =
861
       q)
862
           cindex_k_fold[i,q] <- c_index_tmp$concordance</pre>
863
          }
864
       }
865
       #Store c-index value in a data frame
866
       cindex_k_fold <- as.data.frame(cindex_k_fold)</pre>
867
       #Assign column and row names for c-index table
868
       colnames(cindex_k_fold)<-c("year 1","year 2", "year 3", "year 4", "year 5", "year 6")
       rownames(cindex_k_fold)<-c("fold 1","fold 2", "fold 3", "fold 4", "fold 5")
869
870
       #export results
871
       write.csv(cindex_k_fold, "../deliv/tables/cindex_k_fold.csv", row.names = TRUE)
872
        ```{r K-fold cross-validation analysis stratified by each of the islet AA combinations and continuous
873
874
 covariates using binary groups as discussed in Appendix H Figure 39-73}
875
 #Set a seed value
876
 set.seed(1)
877
 #Apply for-loop to rotate folds for cross-validation
878
 for(i in 1:n){
879
880
 train <- data.frame()</pre>
881
 train inds <-c(1:n)
882
 train_inds <- train_inds[-i]</pre>
883
 test ind <- i
884
 for(j in 1: (n-1)) {train <- rbind(train,splits[[train_inds[j]]])}</pre>
885
 test <- splits[[test ind]]</pre>
886
887
 #Create a covariate list for stratification
888
 strat vars <-
889
 c("GAD65_IAA", "GAD65_ZNT8", "IA2A_ZNT8", "IA2A_IAA_ZNT8", "GAD65_IA2A_IAA_ZNT8",
890
 "A1c_binary", "GLU120_binary")
891
 #Create a list for populating the plot titles
892
 strat_vars_title <- c("GAD65_IAA", "GAD65_ZnT8", "IA-2_ZnT8", "IA-2_IAA_ZnT8", "GAD65_IA-
893
 2_IAA_ZnT8", "HbA1c_binary", "GLU120_binary")
894
895
 #Create a variable with threshold value for continuous covariates
 binary cutoffs <- c("5.25 %","100 mg/dl")
896
897
898
 #Store the number of covariates being used for stratification
899
 n vars <- length(strat vars)</pre>
900
901
 #Apply for loop to rotate folds for cross-validation
902
 for(k in 1:n_vars) {
903
904
 m <- ifelse(k >= 6,k,0)
905
906
 #For the test fold, split the covariate being used for stratification into presence or absence
907
 test_1 <- test %>% filter(.data[[strat_vars[[k]]]] == 1)
908
 test_2 <- test %>% filter(.data[[strat_vars[[k]]]] == 0)
909
910
 #Create "surv" object
```

911 surv obj train <- Surv(train\$T event, train\$status)</pre> 912 913 #Fit model using 'flexsurvreg" function with final multivariate AFT model described in section 914 4.4.2.4 915 fit\_train <- do.call(flexsurvreg, list(formula = surv\_obj\_train ~ GAD65\_IAA + GAD65\_ZNT8 + 916 IA2A\_ZNT8 + IA2A\_IAA\_ZNT8 + GAD65\_IA2A\_IAA\_ZNT8 + HbA1c\_s + log\_GLU120\_s, data = train, 917 dist = "Weibull")) 918 #Check model fit with test fold 919 920 fit\_test\_1 <- survfit(Surv(T\_event, status) ~ 1, data = test\_1)</pre> 921 fit\_test\_2 <- survfit(Surv(T\_event, status) ~ 1, data = test\_2)</pre> 922 surv1 <- summary(fit\_train, newdata = test\_1, type = "survival", B=50, tidy = TRUE) 923 924 surv2 <- summary(fit\_train, newdata = test\_2, type = "survival", B=50, tidy = TRUE) 925 926 varnames <- c("time", "surv", "lower", "upper")</pre> 927 928 fit\_test\_1\_data <- cbind(fit\_test\_1\$time, fit\_test\_1\$surv, fit\_test\_1\$lower, fit\_test\_1\$upper) 929 fit test 1 data <- as.data.frame(fit test 1 data) 930 names(fit\_test\_1\_data) <- varnames</pre> 931 fit\_test\_1\_data\$var <- as.factor(paste(strat\_vars[k], ": 1"))</pre> 932 933 fit\_test\_2\_data <- cbind(fit\_test\_2\$time, fit\_test\_2\$surv, fit\_test\_2\$lower, fit\_test\_2\$upper) 934 fit\_test\_2\_data <- as.data.frame(fit\_test\_2\_data)</pre> 935 names(fit\_test\_2\_data) <- varnames</pre> 936 fit test 2 data\$var <- as.factor(paste(strat vars[k], ": 0")) 937 938 surv 1 avg <- surv1 %>% 939 group by(time) %>% 940 summarise(mean\_est = mean(est, na.rm=TRUE), 941 mean lcl = mean(lcl, na.rm=TRUE), 942 mean ucl = mean(ucl, na.rm=TRUE), 943 var = as.factor(paste(strat\_vars[k], ": 1"))) 944 945 surv 2 avg <- surv2 %>% 946 group\_by(time) %>% 947 summarise(mean est = mean(est, na.rm=TRUE), 948 mean lcl = mean(lcl, na.rm=TRUE), 949 mean\_ucl = mean(ucl, na.rm=TRUE), var = as.factor(paste(strat vars[k], ": 0"))) 950 951 #Generate plots to check goodness-of-fit 952  $if(m != k){$ 953 p < -ggplot() +954 gqtitle(paste("Fold ",i, " Stratified by ", strat vars title[k], sep = "")) + 955  $geom_line(data = surv_1_avg, aes(x = time, y = mean_est, colour = var)) +$ 956 geom\_line(data = surv\_2\_avg, aes(x = time, y = mean\_est, colour = var)) + 957 958  $geom_step(data = fit_test_1_data, aes(x = time, y = surv, colour = var), linetype = 3, size = 1)$ 959 + 960  $geom_step(data = fit_test_2_data, aes(x = time, y = surv, colour = var), linetype = 3, size = 1)$ 961 +

| 962  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 963  | geom_ribbon(data = fit_test_1_data, aes(x = time, ymin = lower, ymax = upper, colour = var, fill                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 964  | = var), linetype = 0, alpha = .2, show.legend = FALSE) +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 965  | geom ribbon(data = fit test 2 data, $aes(x = time, ymin = lower, ymax = upper, colour = var, fill$                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 966  | = var), linetype = 0, alpha = .2, show.legend = FALSE) +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 967  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 968  | geom ribbon(data = surv 1 avg, aes(x = time, ymin = mean lcl, ymax = mean ucl, colour =                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 969  | var, fill = var), linetype = 0, alpha = .2, show.legend = FALSE) +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 970  | geom ribbon(data = surv 2 avg, aes(x = time, vmin = mean lcl, vmax = mean ucl, colour =                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 971  | var, fill = var), linetype = 0, alpha = .2, show.legend = FALSE) +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 972  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 973  | xlab("Time from Derived BL (vears)") +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 974  | vlab("1 - Probability of T1D Diagnosis")                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 975  | ,, · · · · · · · · · · · · · · · ·                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 976  | #View goodness-of-fit plots                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 977  | n                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 978  | P                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 979  | #Export cross-validation plots                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| 980  | agasave( $paste(",/deliv/figures/",i.strat_vars[k],",png", sep = ""), p. width = 16, height = 9, units =$                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 981  | "cm")                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 982  | }                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 983  | 5                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 984  | #Generate plot to check goodness-of-fit                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 985  | if(m = = k)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 986  | n < -aanlot() +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 987  | p < ggproc() '<br>actitle(naste("Fold " i " Stratified by " strat vars title[k] " threshold of " hinary cutoffs[m-5] sen                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 988  | = "")) +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 989  | geom line(data = surv 1 avg. aes(x = time, y = mean est. colour = var)) +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 990  | $geom_ine(data = surv_2 avg_aes(x = time, y = mean_est_colour = var)) +$                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 991  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 992  | geom_step(data = fit_test_1_data, aes(x = time, y = sury, colour = var), linetype = 3, size = 1)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 993  | +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 994  | aeom step(data = fit test 2 data, aes(x = time, y = sury, colour = var), linetype = 3, size = 1)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 995  | +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 996  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 997  | geom ribbon(data = fit test 1 data aes(x = time vmin = lower vmax = upper colour = var fill                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 998  | = var) linetype = 0 alpha = 2 show legend = FALSE) +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 999  | aeom ribbon(data = fit test 2 data aes(x = time vmin = lower vmax = upper colour = var fill                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 1000 | = var) linetype = 0 alpha = 2 show legend = FALSE) +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 1001 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 1002 | geom ribbon(data = surv 1 avg aes(x = time vmin = mean lcl vmax = mean ucl colour =                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 1002 | $y_{ac} = y_{ac} = 0$ alpha = 2 show leagend = FAI SE) +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 1004 | aeom ribbon(data = surv 2 avg aes(x = time vmin = mean lcl vmax = mean ucl colour =                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 1005 | $y_{cont_n} = y_{cont_n} = y_{$ |
| 1006 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 1007 | xlab("Time from Derived BL (vears)") +                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 1008 | $v_{lab}("1 - Probability of T1D Diagnosis")$                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 1009 | Has I Hobabilly of the Diagnosis /                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| 1010 | #View goodness-of-fit plots                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 1011 | n                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| 1012 | ۲                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |

```
1013
 #Export cross-validation plots
1014
 ggsave(paste("../deliv/figures/",i,strat_vars[k],".png", sep = ""), p, width = 16, height = 9, units =
1015
 "cm")
1016
1017
 }
1018
1019
 }
1020
 }
 . . .
1021
         ```{r Cross-validation analysis on pediatric population (age < 12) as discussed in section 4.3.7.3}
1022
1023
        #Set a seed value
1024
        set.seed(1)
1025
        #Assign age threshold of 12
1026
        age_thres <- 12
1027
        #Extract 50% of the pediatric population (age < 12) from the data as test set
1028
        ped inds <- data$IDp[which(data$bAGE < age thres)]</pre>
1029
        ped_inds_test <- sample(ped_inds,round(length(ped_inds)/2), replace = FALSE)</pre>
1030
        #Extract remaining data for model training
1031
        ped inds train <- setdiff(data$IDp,ped inds test)</pre>
1032
        #Prepare train and test data for cross-validation analysis
1033
        train <- data[ped_inds_train,]</pre>
1034
        test <- data[ped inds test,]
1035
        #Create "surv" object
1036
        surv_obj_train <- Surv(train$T_event, train$status)</pre>
1037
        #Fit model using 'flexsurvreg" function - final multivariate AFT model described in section 4.4.2.4
        fit train <- do.call(flexsurvreg, list(formula = surv obj train ~ GAD65 IAA + GAD65 ZNT8 +
1038
1039
        IA2A_ZNT8 + IA2A_IAA_ZNT8 + GAD65_IA2A_IAA_ZNT8 + HbA1c_s + log_GLU120_s, data = train,
1040
        dist = "Weibull"))
1041
1042
        #Test model fit with test data
1043
        fit_test <- survfit(Surv(T_event, status) ~ 1, data = test)</pre>
        surv <- summary(fit_train, newdata = test, type = "survival", B=50, tidy = TRUE)</pre>
1044
1045
1046
        varnames <- c("time", "surv", "lower", "upper")</pre>
1047
1048
        fit_test_data <- cbind(fit_test$time, fit_test$surv, fit_test$lower, fit_test$upper)</pre>
1049
        fit test data <- as.data.frame(fit test data)
1050
        names(fit test data) <- varnames
1051
        surv_avg <- surv %>%
1052
             group by(time) %>%
1053
             summarise(mean est = mean(est, na.rm=TRUE),
1054
                   mean_lcl = mean(lcl, na.rm=TRUE),
1055
                   mean_ucl = mean(ucl, na.rm=TRUE),
1056
                   )
1057
        #Generate goodness-of-fit plot
1058
        p < -ggplot() +
1059
        gqtitle(paste("Cross validation on pediatric population: Age < ",age thres, sep = "")) +
1060
        geom_line(data = surv_avg, aes(x = time, y = mean_est)) +
1061
        geom_step(data = fit_test_data, aes(x = time, y = surv), linetype = 3, size = 1) +
1062
        geom_ribbon(data = fit_test_data, aes(x = time, ymin = lower, ymax = upper), linetype = 0, alpha
1063
        = .2, show.legend = FALSE) +
```

geom_ribbon(data = surv_avg, aes(x = time, ymin = mean_lcl, ymax = mean_ucl), linetype = 0, 1064 1065 alpha = .2, show.legend = FALSE) + 1066 #Add x and y labels 1067 xlab("Time from Derived Baseline (years)") + 1068 ylab("1 - Probability of T1D Diagnosis") 1069 *#view plot* 1070 р 1071 #Export plot to "Figures" folder 1072 ggsave(paste("../deliv/figures/ped_validation_",age_thres,"c.png", sep = ""), p, width = 16, height = 1073 9, units = "cm") 1074 1075 ```{r Cross-validation analysis on pediatric population (age < 12) as discussed in seciton 4.3.7.3 - C-1076 1077 index table} 1078 #Assign maximum year for c-index calculation 1079 yrs for cindex <- 6 1080 #Create a matrix to store c-index values 1081 cindex_peds <- matrix(NA,nrow = 1, ncol = yrs_for_cindex)</pre> 1082 #Use "survreg" with the final multivariate AFT model described in section 4.4.2.4 to compute c-index fit_train_concordance<- survreg(Surv(T_event, status) ~ GAD65_IAA + GAD65_ZNT8 + IA2A_ZNT8 + 1083 1084 IA2A_IAA_ZNT8 + GAD65_IA2A_IAA_ZNT8 + HbA1c_s +log_GLU120_s , data = train ,dist = 1085 "weibull") 1086 #Compute c-index till six years with one-year increments 1087 for(q in 1:yrs_for_cindex){ 1088 c_index_tmp <- concordance(object = fit_train_concordance, newdata = test, ymin = 0,ymax = 1089 q) 1090 cindex_peds[1,q] <- c_index_tmp\$concordance 1091 } 1092 #Store the c-index values in a data frame 1093 cindex peds <- as.data.frame(cindex peds)</pre> 1094 #Create columns and rows names for c-index table colnames(cindex_peds)<-c("year 1","year 2", "year 3", "year 4", "year 5", "year 6") 1095 1096 rownames(cindex_peds)<-c("Peds c-index")</pre> 1097 #Export the c-index table 1098 write.csv(cindex_peds, "../deliv/tables/cindex_peds.csv", row.names = TRUE) 1099 ```{r Model performance using time dependent ROC as discussed in section 4.3.7.1} 1100 1101 #Select data for time dependent ROC analysis and convert status to 0 and 1 to use predict function 1102 data_for_ROC<-data %>% 1103 select(IDp,T event,status,GAD65 IAA,GAD65 ZNT8, IA2A ZNT8, IA2A IAA ZNT8, 1104 GAD65_IA2A_IAA_ZNT8 , log_GLU0_s ,HbA1c_s ,log_GLU120_s) %>% 1105 *mutate(status=status-1)* 1106 #Identify missing covariate value 1107 *aa=which(complete.cases(data for ROC)==F)* 1108 #Fit the model using the "survreg" function 1109 fit_weib<- survreg(Surv(T_event, status) ~ GAD65_IAA + GAD65_ZNT8 + IA2A_ZNT8 + 1110 IA2A IAA ZNT8 + GAD65 IA2A IAA ZNT8 + HbA1c s +log GLU120 s , data = data for ROC, dist = 1111 "weibull") 1112 #Extract the linear predictor 1113 data_for_ROC\$lp <- predict(fit_weib, type = "lp")</pre> 1114 #Define a helper function to evaluate at various time points

```
1115
        survivalROC helper <- function(t) {</pre>
         survivalROC(Stime
                                 = data_for_ROC$T_event,
1116
1117
                 status
                            = data_for_ROC$status,
1118
                 marker
                             = data_for_ROC$lp,
1119
                 predict.time = t,
1120
                              = "KM")#,span = 0.25 * nrow(data_for_ROC)^(-0.20))
                 method
1121
        }
1122
        #Evaluate every 0.5 years
1123
        survivalROC_data <- tibble(t =seq(0.5,5.5,by=1)) %>%
1124
         mutate(survivalROC = map(t, survivalROC_helper),
1125
              ## Extract scalar AUC
1126
              auc = map_dbl(survivalROC, magrittr::extract2, "AUC"),
1127
              ## Put cut off dependent values in a data frame
1128
              df_survivalROC = map(survivalROC, function(obj) {
1129
               as_data_frame(obj[c("cut.values","TP","FP")])
1130
              }))%>%
1131
         dplyr::select(-survivalROC) %>%
1132
         unnest() %>%
1133
         arrange(t, FP, TP) %>%
         mutate(FP=1-FP,TP=1-TP,auc=1-auc)
1134
1135
        #Generate ROC curves
        p ROC <-qqplot(data = survivalROC data, mapping = aes(x = FP, y = TP)) +
1136
1137
           ggtitle("Time dependent ROC analysis: model predictions on full analysis set")+
1138
           geom_point() +
1139
           geom_line() +
           geom label(data = survivalROC data %>% dplyr::select(t,auc) %>% unique,
1140
1141
                  mapping = aes(label = sprintf("%.3f", auc)), x = 0.5, y = 0.5) +
           facet_wrap( ~ t, labeller = labeller(t = c("0.5" = "0.5 \text{ years}", "1.5" = "1.5 \text{ years}", "2.5" = "2.5")
1142
        years", "3.5" = "3.5 years", "4.5" = "4.5 years", "5.5" = "5.5 years"))) +
1143
1144
           xlab("FPR")+
1145
           ylab("TPR") +
1146
           theme bw() +
1147
           theme(axis.text.x = element text(angle = 90, viust = 0.5),
1148
               legend.key = element blank(),
1149
               plot.title = element_text(hjust = 0.5),
1150
               strip.background = element_blank())
1151
        #View ROC curves
1152
        p ROC
1153
        #Export plot
        ggsave(paste("../deliv/figures/survival ROC.png", sep = ""),p ROC, width = 16, height = 13, units =
1154
1155
        "cm")
         . . .
1156
         ```{r External validation using DAISY dataset as discussion in section 4.3.7.4}
1157
1158
 #Create a "surv" object
1159
 surv_obj_train <- Surv(data$T_event, data$status)</pre>
1160
 #Train model - final multivariate AFT model described in section 4.4.2.4
 fit train <- do.call(flexsurvreg, list(formula = surv obj train ~ GAD65 IAA + GAD65 ZNT8 +
1161
 IA2A_ZNT8 + IA2A_IAA_ZNT8 + GAD65_IA2A_IAA_ZNT8 +HbA1c_s +log_GLU120_s, data = data,
1162
 dist = "Weibull"))
1163
1164
 #Test the model with external data from DAISY study
1165
 test <- data_daisy
```

1166 fit\_test <- survfit(Surv(T\_event, status) ~ 1, data = test)</pre> 1167  $surv < -summary(fit_train, newdata = test, type = "survival", B=50, tidy = TRUE)$ 1168 varnames <- c("time", "surv", "lower", "upper")</pre> 1169 fit\_test\_data <- cbind(fit\_test\$time, fit\_test\$surv, fit\_test\$lower, fit\_test\$upper)</pre> 1170 fit\_test\_data <- as.data.frame(fit\_test\_data)</pre> names(fit\_test\_data) <- varnames</pre> 1171 1172 surv\_avg <- surv %>% 1173 group\_by(time) %>% 1174 summarise(mean\_est = mean(est, na.rm=TRUE), 1175 mean\_lcl = mean(lcl, na.rm=TRUE), 1176 mean\_ucl = mean(ucl, na.rm=TRUE)) 1177 #Generate plot to check goodness-of-fit p < -ggplot() +1178 1179 ggtitle("External Validation using DAISY dataset") + 1180  $geom_line(data = surv_avg, aes(x = time, y = mean_est)) +$ 1181  $geom_step(data = fit_test_data, aes(x = time, y = surv), linetype = 3, size = 1) +$ 1182 geom\_ribbon(data = fit\_test\_data, aes(x = time, ymin = lower, ymax = upper ), linetype = 0, 1183 alpha = .2, show.legend = FALSE) + 1184 geom ribbon(data = surv avg, aes(x = time, ymin = mean lcl, ymax = mean ucl), linetype = 0,1185 alpha = .2, show.legend = FALSE) +1186 xlab("Time from Derived BL (years)") + 1187 ylab("1 - Probability of T1D Diagnosis") 1188 #View goodness-of-fit 1189 р 1190 1191 #Export cross-validation plot 1192 ggsave(paste("../deliv/figures/Daisy\_External\_Validation.png", sep = ""), p, width = 16, height = 9, units = "cm") 1193 1194 #Assign maximum year for c-index calculation 1195 vrs for cindex <- 6 1196 #Create a matrix to store c-index values cindex\_daisy <- matrix(NA,nrow = 1, ncol = yrs\_for\_cindex)</pre> 1197 1198 1199 #Use "survreg" with the final multivariate AFT model described in section 4.4.2.4 to compute c-index 1200 fit\_train\_concordance<- survreg(Surv(T\_event, status) ~ GAD65\_IAA + GAD65\_ZNT8 + IA2A\_ZNT8 + 1201 IA2A\_IAA\_ZNT8 + GAD65\_IA2A\_IAA\_ZNT8 + HbA1c\_s +log\_GLU120\_s , data = data ,dist = 1202 "weibull") 1203 #Compute c-index till six years with one-year increments 1204 for(q in 1:yrs\_for\_cindex){ 1205 c index tmp <- concordance(object = fit train concordance, newdata = data daisy, ymin = 1206 0,ymax = q1207 cindex\_daisy[1,q] <- c\_index\_tmp\$concordance 1208 } 1209 #Store the c-index values in a data frame 1210 cindex\_daisy <- as.data.frame(cindex\_daisy)</pre> 1211 #Create columns and rows names for c-index table colnames(cindex\_daisy)<-c("year 1","year 2", "year 3", "year 4", "year 5", "year 6") 1212 1213 rownames(cindex\_daisy)<-c("Daisy c-index")</pre> #Export the c-index table 1214 1215 write.csv(cindex\_daisy, "../deliv/tables/cindex\_daisy.csv", row.names = TRUE) 1216