Modeling the Progression of Type-I Diabetes using Machine Learning

Manya Chadha, Elizabeth Holden, Sarah Nguyen

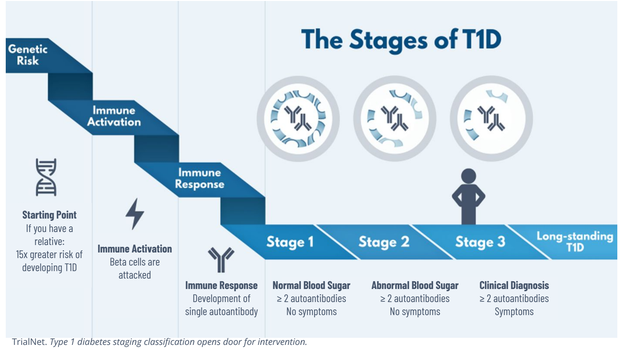
Benaroya Research Institute

Background

Type 1 Diabetes (T1D) is an autoimmune disease characterized by the destruction of insulin-producing beta cells. Insulin is a hormone essential for maintaining blood sugar, without which patients develop excessive thirst, urination, and weight loss and can suffer life threatening complications. The presence of **islet autoantibodies (AABs)** is the strongest known predictor of T1D development.

Why early prediction matters:

- Individuals with **2+ AABs** will eventually develop T1D.
- **20% of those with 1 AAB** develop 2+ AABs within 5 years.
- Early prediction enables **treatment to delayed onset** and **reduces severe complications** like diabetic ketoacidosis (DKA).
- **DKA rates at diagnosis exceed 30%**, but early monitoring can **reduce this 10-fold** in clinical trials.



TrialNet. *Type 1 diabetes staging classification opens door for intervention.*

Project Goals

Validate an Existing Model

- The **Benaroya Research Institute (BRI)** developed a Bayesian survival model for **T1D risk prediction**[1].
- We aim to **test how the model generalizes** across two external datasets: **FRIDA and DPT-1**.

Explore Machine Learning Alternatives

- Can machine learning (ML) models improve **accuracy & explainability** over the Bayesian approach?
- Evaluate **Random Survival Forest (RSF)**, **dynforest**, and **Joint Modeling** as alternative ML approaches.

Build a Web Application

- Integrate models into a **clinician-friendly web tool** for **real-time risk prediction** & **personalized monitoring.**

[1]Pribitzer et al. *"Beyond Stages: Predicting Individual Time Dependent Risk for Type 1 Diabetes." The Journal of clinical endocrinology and metabolism.*
[2] Qi et al. (2023). An Effective Meaningful Way to Evaluate Survival Models. *Proceedings of the 40th International Conference on Machine Learning* (PMLR 202), 28295–28318.

Datasets

Dataset	Age range	# of unique patients	# of observations	Details
TrialNet Pathway to Prevention	2 - 45	6193	33,821	Screening study of individuals at risk for T1D, based in the US
FRIDA	2 - 10	420	2,277	T1D screening study based in Germany
DPT-1	0 - 50	274	1,050	Insulin therapy trial based in the US

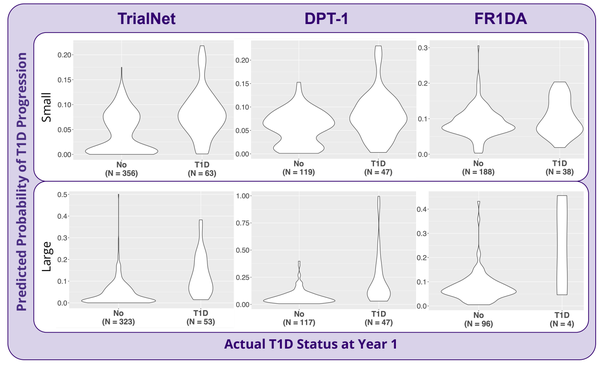
Longitudinal data

BRI Validation

Parameter	Small	Medium	Large
Age	✓	✓	✓
Auto antibody status	✓	✓	✓
HbA1c	✓	✓	✓
Fasting Glucose	✓	✓	✓
2 hour Glucose		✓	✓
Early C-Peptide Change			✓
C-Peptide AUC			✓

Area Under the Curve Results

Model	Landmark T =1			Landmark T = 3		
	Trial Net	DPT-1	FRIDA	Trial Net	DPT-1	FRIDA
Small	0.73	0.66	0.56	0.71	0.62	0.59
Med.	0.73	0.71	0.77	0.76	0.66	0.60
Large	0.82	0.80	0.76	0.83	0.77	No data



Methodology

The Role of Survival Modeling

What is Survival Modeling?

A technique used for **time-to-event prediction** such as analyzing the duration of time until a patient is diagnosed with T1D.

- **Event:** Onset of Type 1 diabetes.
- **Time:** Time from the beginning of an observation period to an event, end of study, or loss of contact from study subject.
- **Censored observations**: Occurs when a patient does not experience the event of interest within the duration of the study, i.e. due to drop out or end of study period.

Modeling Approaches

We explore multiple options that extend survival modeling:

Random Survival Forest (RSF)
- An extension of random forest used to handle right censored survival data and give time-to-event predictions.
- Does not require explicit specification of covariate relationships, thus can capture nonlinear interactions.

Dynamic Forest (DynForest)
- Dynamic extension of RSF, allowing for **time-fixed** and **time-dependent** covariates.
- Captures **longitudinal biomarker evolution** using linear mixed models.

Joint Survival Modeling
- **Combines longitudinal and survival sub-models** through shared random effects to predict T1D risk.
- Captures biomarker trajectories through linear modelling.
- Dynamically updates based on historical data.

Evaluation

We evaluate the models using two different methods:

Landmarked Area Under the Curve (AUC-ROC)

- AUC measures **discriminative ability**, i.e., how well the model distinguishes between patients who will and won't develop T1D within a "landmarked" time frame.
- **Landmarking (LM)**: Go "landmark" time points back from last observation and generate predictions for that "landmarked" interval, using past data.

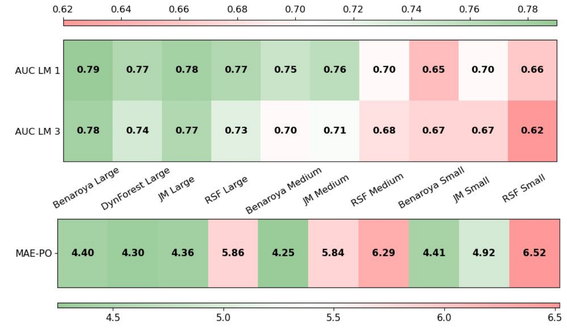
Mean Absolute Error (MAE) - PO[2]

- Calculates pseudo-observations (derived from **Kaplan-Meier estimates**) and applies re-weighting to compute MAE.
- Why MAE-PO? Measures how closely predicted T1D onset aligns with actual outcomes even with censored data.

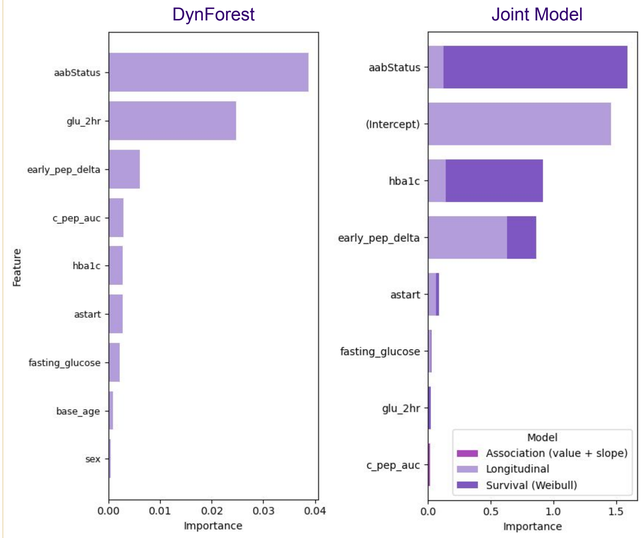
$$e_{\text{pseudo-obs}}(t_i, \mathcal{D}) = N \times \hat{\theta} - (N-1) \times \hat{\theta}^{-i} \qquad \text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\boldsymbol{x}_i - \boldsymbol{x}|$$

Results & Discussion

Landmarked AUC and MAE-PO



Feature Importance



BRI model validation

- The BRI AUC values for FR1DA & DPT-1 indicates that the **small Benaroya model doesn't generalize well to new patients.**
- The BRI large model, however, performs well on external validation data. This indicates that when 2-hour glucose, early peptide delta, and c-peptide AUC are included, the **large model** is able to **determine T1D status more accurately on unseen data** than the other BRI models.

BRI model vs. Machine Learning models

Model Size	Small	Medium	Large
Best AUC	Joint Model, 0.69	Joint Model, 0.74	Benaroya, 0.78
Best MAE	Benaroya, 4.41	Benaroya, 4.25	DynForest, 4.30

- The **best overall models**, balancing AUC and MAE-PO, **are Benaroya Large** and **Joint Model Large**.
- Among ML models, joint modeling is the most competitive alternative, offering strong AUC and reasonable MAE-PO, making it a viable ML alternative for improved accuracy and explainability.
- An **added advantage of the joint model** is the ability to **dynamically update predictions as historical patient data becomes available** through electronic health records.
- The feature importance graphs reiterate AAB Status is the strongest predictor of T1D development, across different models.

Web Application

BRI and ML models are now accessible through a single web page using Dash, with each model being served as a separate API endpoint built via Plumber API.

Acknowledgements: We thank Drs. Hannah DeBerg, Stephan Pribitzer, Carla Greenbaum, and Cate Speake of the Benaroya Research Institute for their invaluable support and guidance on this project.