TEMPORAL FEATURE-DRIVEN TECHNIQUE FOR PRECISION MEDICINE: EARLY IDENTIFICATION OF INDIVIDUALS AT HIGH RISK FOR STAGE 3 TYPE 1 DIABETES

Shahrukh Iqbal[1], Reenal Boddul[1] Åke Lernmark[2], Mikael Nordenstjerna[1], Paul Pettersson[1,3]
[1]Mainly AI, Stockholm, Sweden. [2]Department of Clinical Sciences, Lund University CRC, Skåne University Hospital, Malmö, Sweden, [2]Mälardalens University, Sweden

BACKGROUND

Early identification of individuals at high risk of developing Type 1 diabetes (T1D) is essential for preventing or delaying the clinical onset of the disease. We propose a novel feature-driven technique to identify high-risk individuals for T1D.

METHODS

We selected 7100 participants (291 developed T1D) over the age between 365 days and 5159 days and more than 5 visits, from the TEDDY cohort [1]. Genetic, demographic, and physical features were considered for analysis. A probabilistic risk-score which is a cumulative, non-decreasing and bounded function of 'autoantibody persistence over time' is proposed (**Figure 1**). The risk-score was computed for T1D-positive and T1D-negative participants using bootstrap-resampling (**Table-1**). A time-to-event analysis was performed using risk-score and dataset-specific features. The baseline model was Cox-Proportion Hazard (CPH) Model, which was compared with two non-linear models: Decision Tree (DT) Survival Model, and Random Survival Forest (RSF). Uno's Concordance Index (CI), Mean Temporal Area-under-the-curve (AUC(t)) and Integrated Brier Score (IBS) were used for evaluation.

Train-Test split was 80-20, and Monte-Carlo cross validation *(n=100)* (MCCV) was used for evaluation (**Table-2**).

$$Score(x^{ab}, T) = \sum_{t=1}^{t=T} x_t^{ab} \cdot f(x_t^{ab}) \cdot f(t) \; where,$$

$$ab \in \{GAD, MIAA, IA2A, ZnT8A, 2^+\}$$

$$x_t^{ab} = \begin{cases} 0, & \text{if ab is not persistent at time t} \\ 1, & \text{if ab is persistent at time t} \end{cases}$$

$$f(x_t^{ab}) = \mathbf{Pr}(x_t^{ab} | x_{t-1}^{ab}) \text{ i.e., conditional persistence probability at t}$$

$$f(t) = \mathbf{Pr}(t_{previous} \leq t \leq t_{current}) \text{ i.e., Likelihood of age between visits}$$

Figure 1. *Score(x^{ab},T)* is the risk-score for individual ***x*** at time ***T*** for autoantibody ***ab***. Time is measured in days windows. $2^+ = 2$ or more autoantibodies were observed.

CONCLUSION

This study proposes a novel feature-driven technique to differentiate high-risk population for Stage 3 type 1 diabetes. The technique is based on treating autoantibody persistence as a temporal phenomenon and the development of probabilistic risk-score that accounts for the variability in autoantibody persistence. The model developed in this study can be used for for screening and early detection of Stage 3 type diabetes.

RESULTS

	T1D-Negative	T1D-Positive
Autoantibody (AAB)	Mean-Score $\pm$ variance	Mean-Score $\pm$ variance
GADA	$0.007_{\pm 0.049}$	$\mathbf{0.124_{\pm 0.161}}$
IAA	$0.0023_{\pm 0.024}$	$\mathbf{0.054_{\pm 0.105}}$
IA-2A	$0.0013_{\pm 0.016}$	$\mathbf{0.036_{\pm 0.071}}$
ZnT8A	$0.001_{\pm 0.012}$	$\mathbf{0.022_{\pm 0.048}}$
Multi-AAB	$0.0019_{\pm 0.019}$	$\mathbf{0.051_{\pm 0.091}}$

Table-1: Mean Risk-Score ± variance for T1D-Negative and T1D-Positive individuals. The Mean risk-score is computed using bootstrap resampling (n= 100,000).

Metric \ Model	CPH	DT	RSF
CI	$0.975_{\pm 0.015}$	$0.916_{\pm 0.024}$	$\mathbf{0.988}_{\pm 0.003}$
AUC(t)	$0.988_{\pm 0.009}$	$0.954_{\pm 0.014}$	$\mathbf{0.994}_{\pm 0.002}$
IBS	$0.336_{\pm 0.123}$	$0.012_{\pm 0.002}$	$\mathbf{0.009}_{\pm 0.001}$

Table-2: Model performance of CPH, DT and RSF. We report Mean Performance $\pm$ standard error estimated using Monte Carlo Cross Validation (n=100) across test-set.

ACKNOWLEDGEMENT

This work has been performed within ASSET innovation milieu and supported by VINNOVA, the Swedish Governmental Innovation Agency (Grant number 2021-02 684)

[1] Hagopian et al. "TEDDY–The Environmental Determinants of Diabetes in the Young: an observational clinical trial." *Annals of the New York Academy of Sciences* 1079.1 (2006): 320-326.

Contact

Mainly.AI