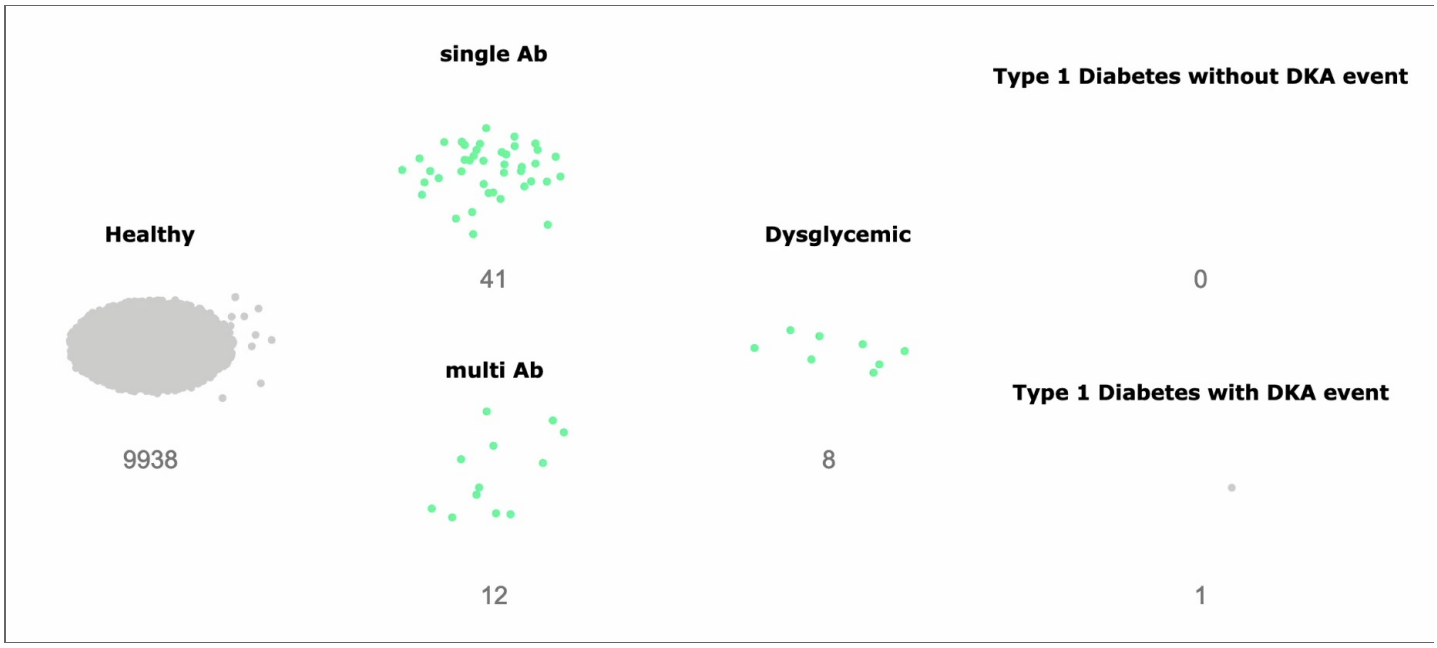
Beyond Predictive Powers

A Comprehensive Comparison of T1D Risk Models

Lauric Ferrat

Introduction



Simulation of T1D screening by Gonçalo Lieria

What is a good predictive model?

1. Accurate
2. Cost is not prohibitive
3. Patient burden is low
4. Accessible

Method

- Generate a large set of predictive models;
- Assess them not only in term of predictive power but in term of cost and patient time.

Variables linked to T1D risk

1. Clinical
   - Age, logarithm of age
   - BMI, Z-BMI
   - Sex
2. Immunology
   - IA2A (Positive or Negative)
   - Autoantibody group (GAD, IAA, IA2A, GAD-IAA, GAD-IA2A, IAA-IA2A, GAD-IAA-IA2A)
3. Genetic
   - GRS2
4. Metabolic
   - derived measured from Oral Glucose Tolerance Test : AUC glucose, AUC C-peptide, C peptide index 30, β2-score, $Index_{60}$ [1], DPTRS[2] and $DPTRS_{60}$ [2], M_{120} [3], CPH[4] and LR[4]
   - HbA1c

1: Sosenko, J. M. et al. A New Approach for Diagnosing Type 1 Diabetes in Autoantibody-Positive Individuals Based on Prediction and Natural History. Diabetes Care 38, 271–276 (2015).
2: Simmons KM et al. One-Hour Oral Glucose Tolerance Tests for the Prediction and Diagnostic Surveillance of Type 1 Diabetes. J Clin Endocrinol Metab. 2020 Nov 1;105(11):e4094–101. doi: 10.1210/clinem/dgaa592. PMID: 32844178; PMCID: PMC7514797.
3: Bediaga NG et al. Simplifying prediction of disease progression in pre-symptomatic type 1 diabetes using a single blood sample. Diabetologia. 2021 Nov;64(11):2432-2444. doi: 10.1007/s00125-021-05523-2. Epub 2021 Aug 2. PMID: 34338806; PMCID: PMC8494707.
4: Weiss et al. Progression likelihood score identifies substages of presymptomatic type 1 diabetes in childhood public health screening. Diabetologia 65, 2121–2131 (2022). https://doi.org/10.1007/s00125-022-05780-9

Model fitting – generation of combinations of variables

> 2 millions different combinations of variables

Some rules when added to only generate plausible models such as

- **BMI** and **z-BMI** would not appear in the same model;
- A model would not have two OGTT derived variables simultaneously.

Here a list of the 100 first ones (out of 2154 models):

	formula
1	GRS2
2	age
3	BMI
4	log(age)
5	Index60
6	AUC-glucose
7	AUC-ceptide
8	z-BMI
9	C-peptide-index-30
10	beta2-score
11	Hba1c
12	GRS2 + IA2
13	GRS2 + AB-group
14	GRS2 + Sex
15	GRS2 + age
16	IA2 + age
17	AB-group + age
18	Sex + age
19	GRS2 + BMI
20	IA2 + BMI
21	AB-group + BMI
22	Sex + BMI
23	age + BMI
24	GRS2 + log(age)
25	IA2 + log(age)
26	AB-group + log(age)
27	Sex + log(age)
28	BMI + log(age)
29	GRS2 + Index60
30	IA2 + Index60
31	AB-group + Index60
32	Sex + Index60
33	age + Index60
34	BMI + Index60

Model fitting

The TrialNet Pathway to Prevention dataset was split in two to train and test the predictive model

	Training dataset (N=1524)	Testing dataset (N=1551)	P-value
Sex			0.006
Female	728 (47.8%)	663 (42.7%)	
Male	796 (52.2%)	888 (57.3%)	
age	9.22 (4.3)	9.15 (4.2)	0.628
GRS2	13.24 (2.1)	13.44 (2.1)	0.01
IA2			<0.001
IA2	621 (40.7%)	731 (47.1%)	
IA2 free	903 (59.3%)	820 (52.9%)	
BMI	18.96 (4.5)	18.61 (4.6)	0.032
AUC Cpeptide	20.11 (9.5)	20.80 (10.2)	0.052
AUC glucose	522.83 (103.4)	542.21 (107.8)	<0.001
Hba1c	5.09 (0.3)	5.11 (0.3)	0.26
Stage			<0.001
0	760 (49.9%)	575 (37.1%)	
1	543 (35.6%)	669 (43.1%)	
2	221 (14.5%)	307 (19.8%)	
T1D			<0.001
T1D	631 (41.4%)	515 (33.2%)	
T1D free	893 (58.6%)	1036 (66.8%)	

Model Cost

Cost of a prediction[1] = clinician time + secretary time + lab test + patient time + travel

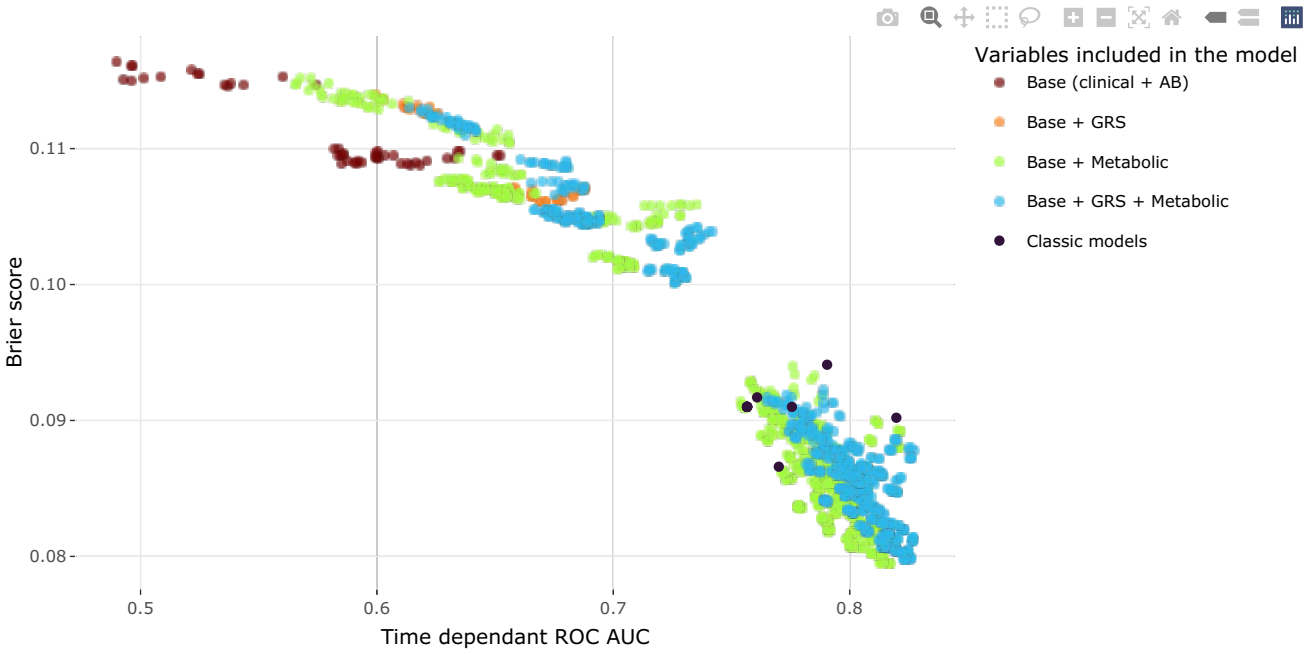
Direct costs were calculated using Medicare reimbursement rates[2].

Indirect costs were calculated using wage data from the US Bureau of Labor Statistics[3].

1. Ping Zhang et al. Costs of Screening for Pre-diabetes Among U.S. Adults: A comparison of different screening strategies. Diabetes Care 1 September 2003; 26 (9): 2536–2542. https://doi.org/10.2337/diacare.26.9.2536
2. Physician fee schedule. Centers for Medicare & Medicaid Services website. 2023. https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeeSched/index.html?redirect=/PhysicianFeeSched/. Accessed June, 2023
3. US Department of Labor. Occupational employment statistics. Bureau of Labor Statistics website. 2023. http://bls.gov/oes/. Accessed June 30, 2023

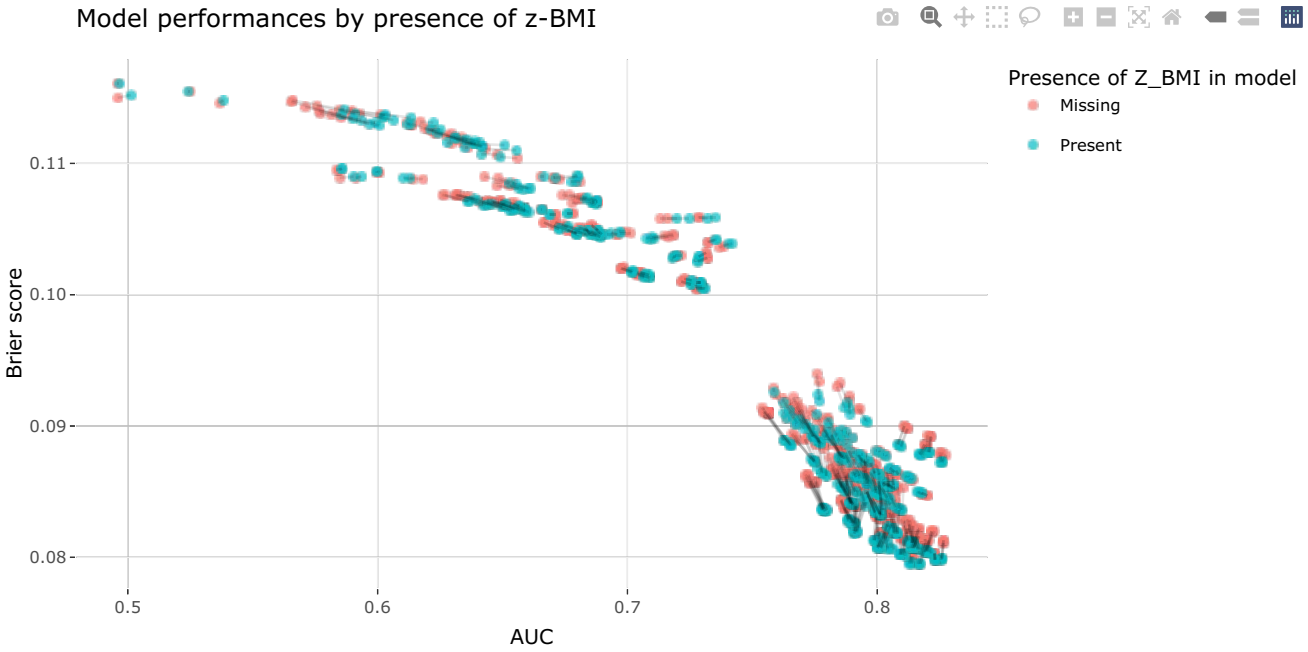
Results

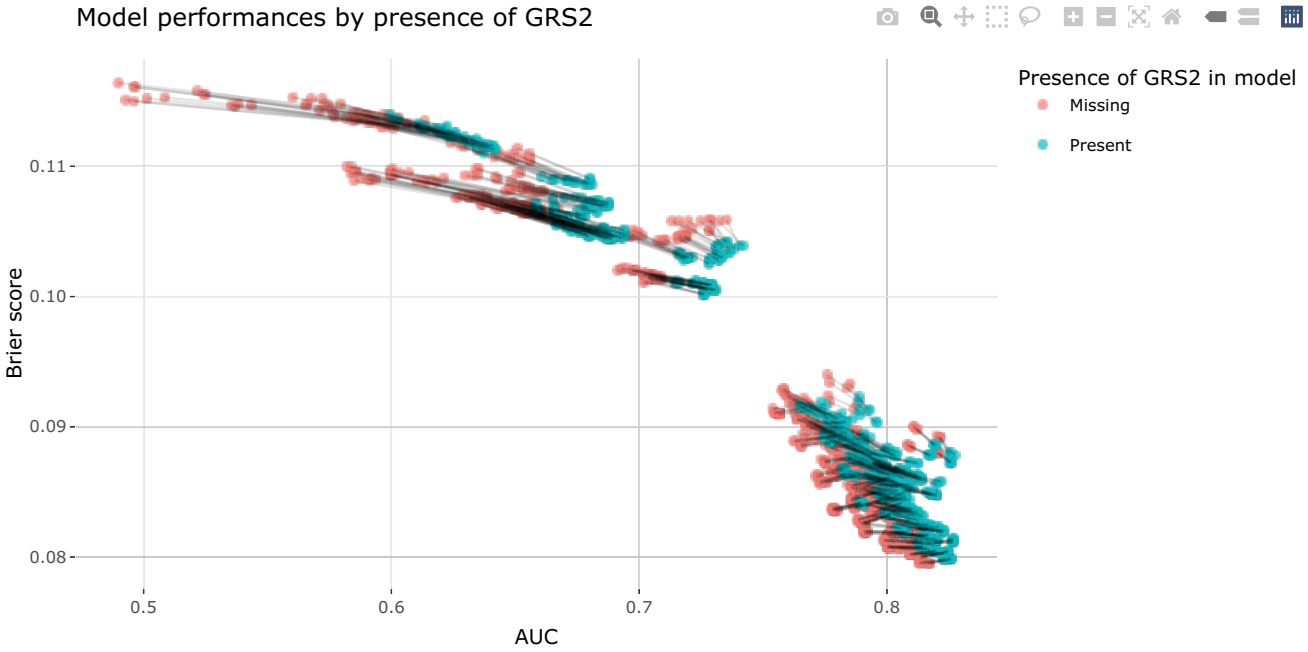
Model performance in single AB

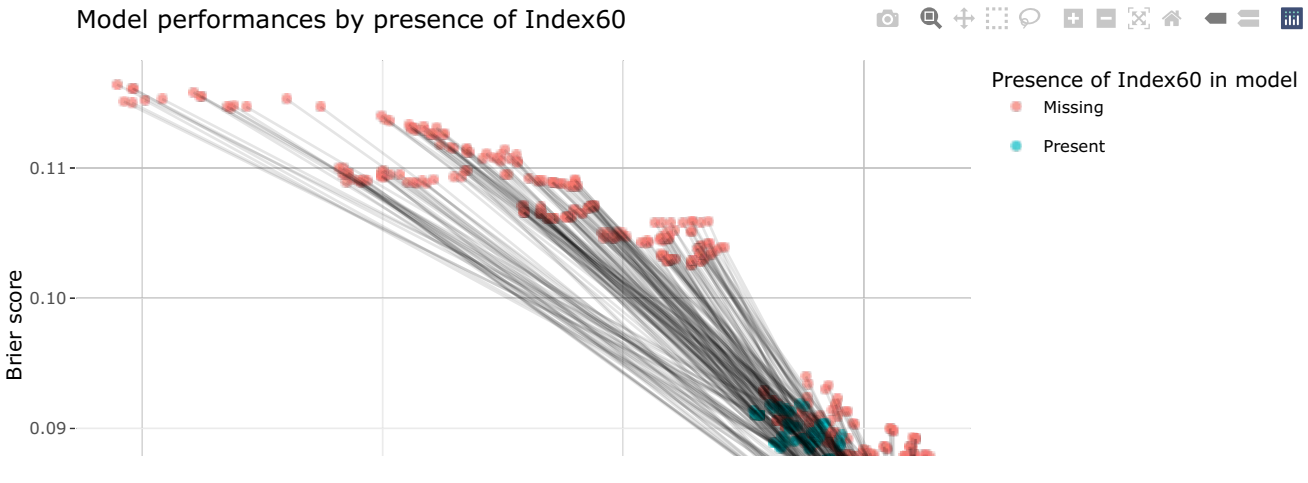


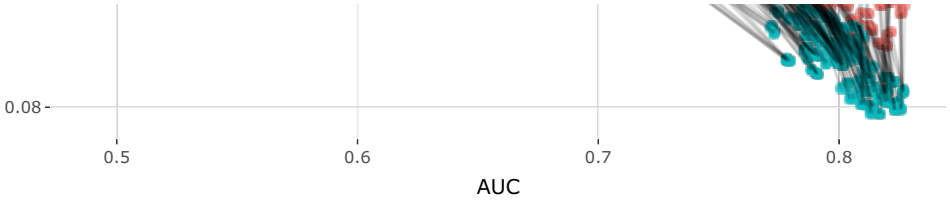
Variable importance in single AB

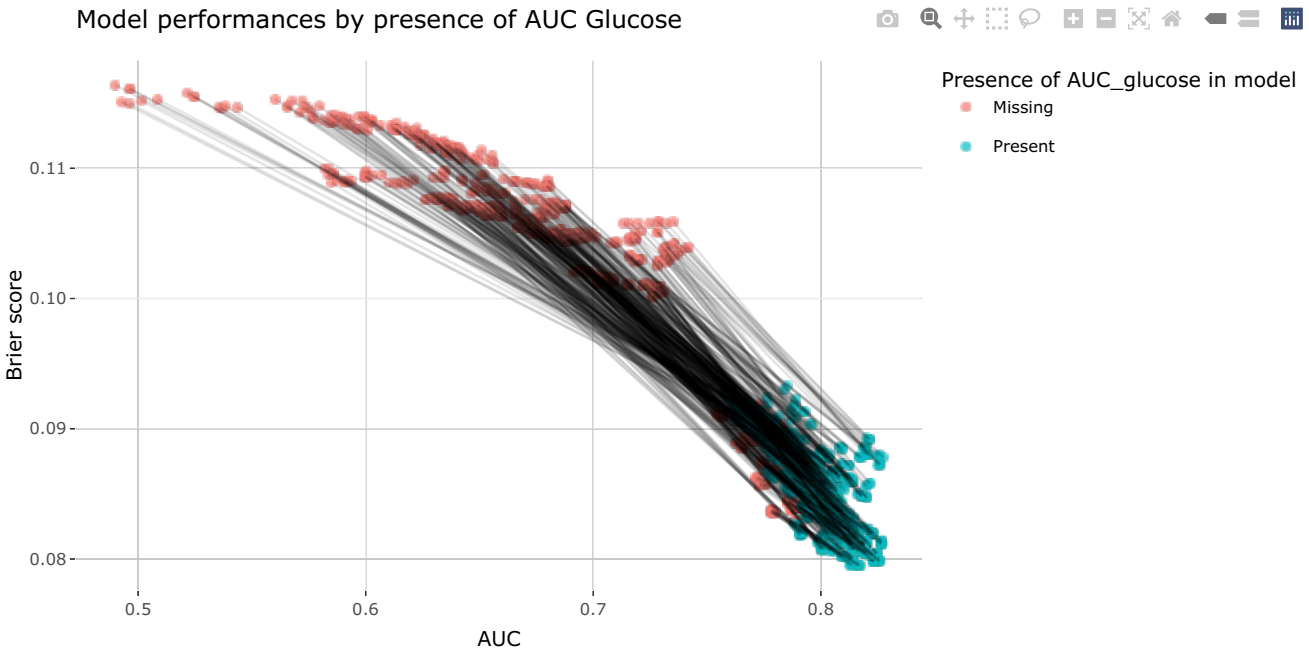
- z-BMI
- GRS2
- Index 60
- AUC glucose



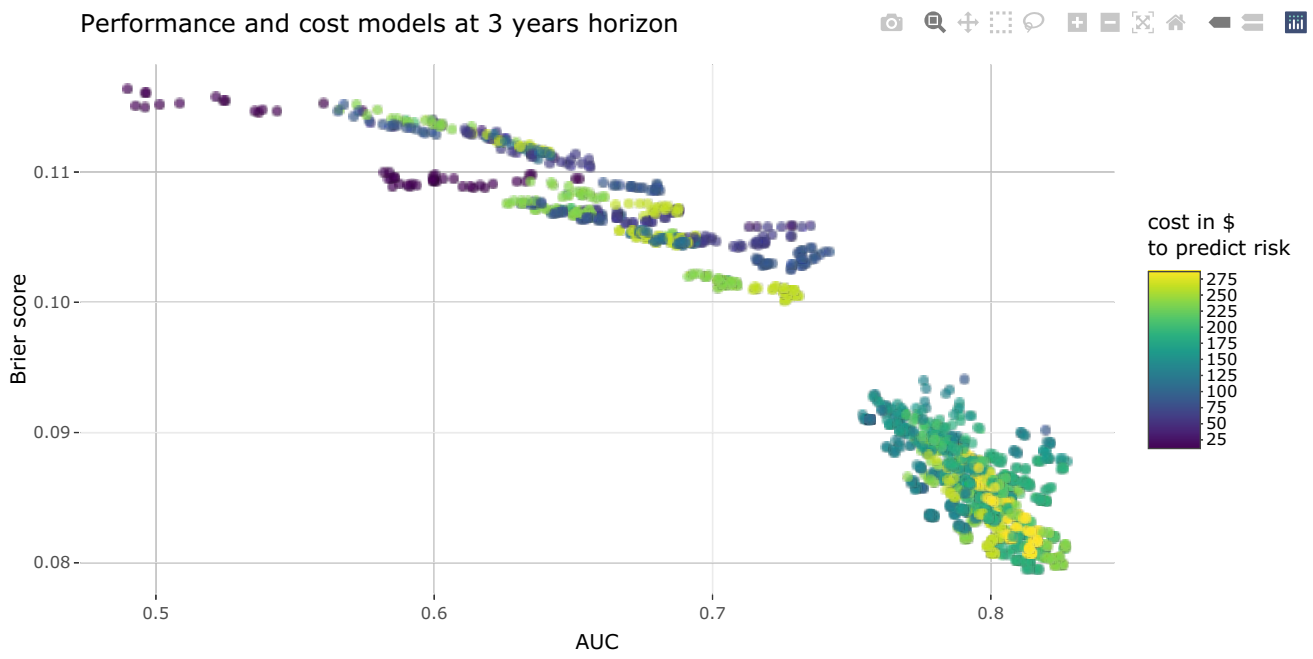




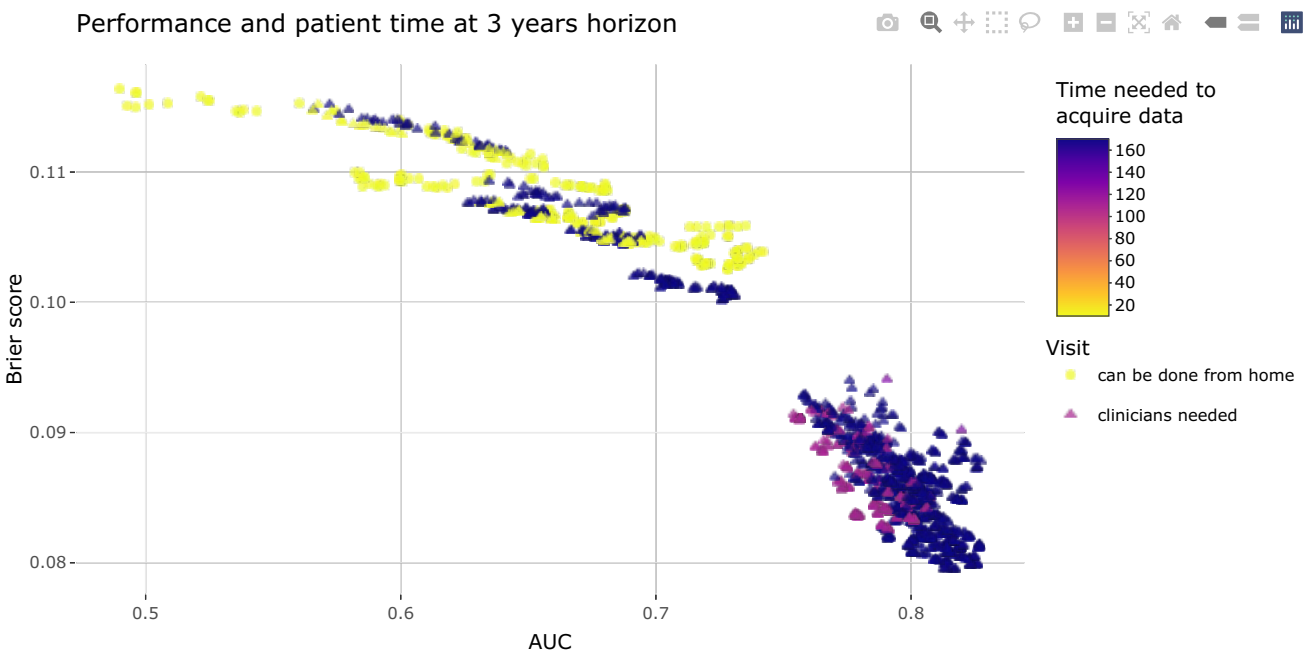

0.08
0.5
0.6
0.7
0.8
AUC


Model performances by presence of AUC Glucose
Presence of AUC_glucose in model
Missing
Present
Brier score
0.11
0.10
0.09
0.08
0.5
0.6
0.7
0.8
AUC

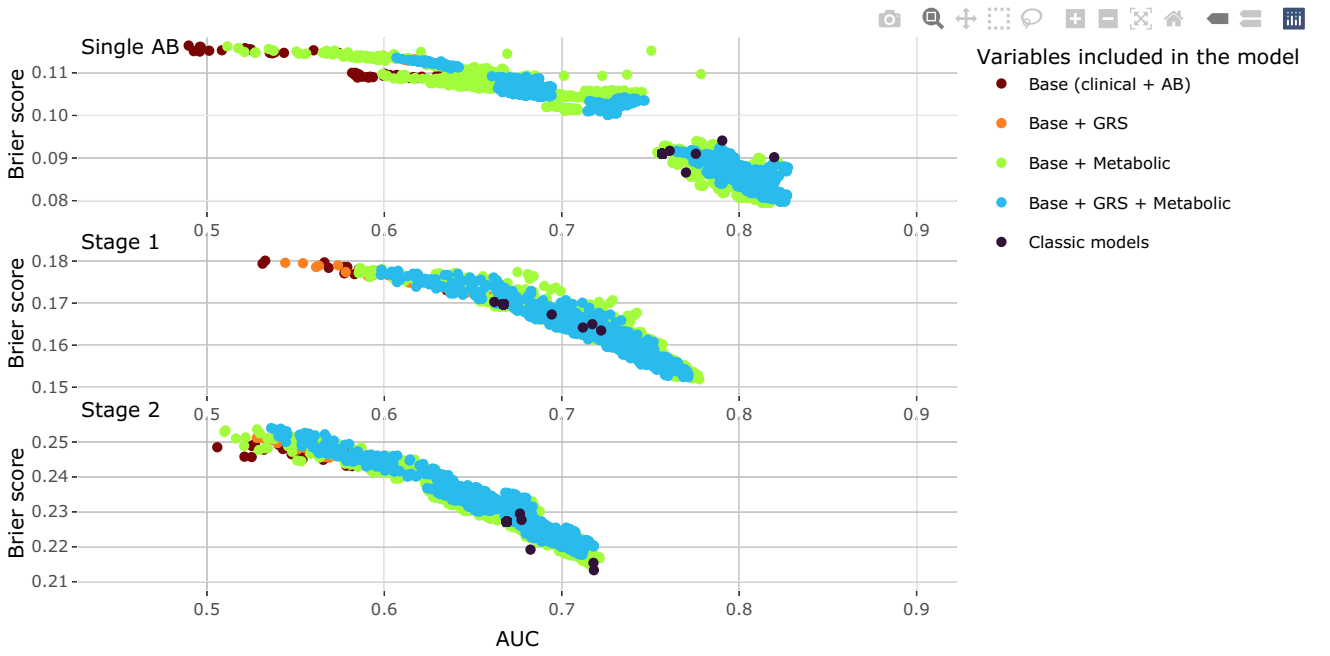
Model Cost at single AB



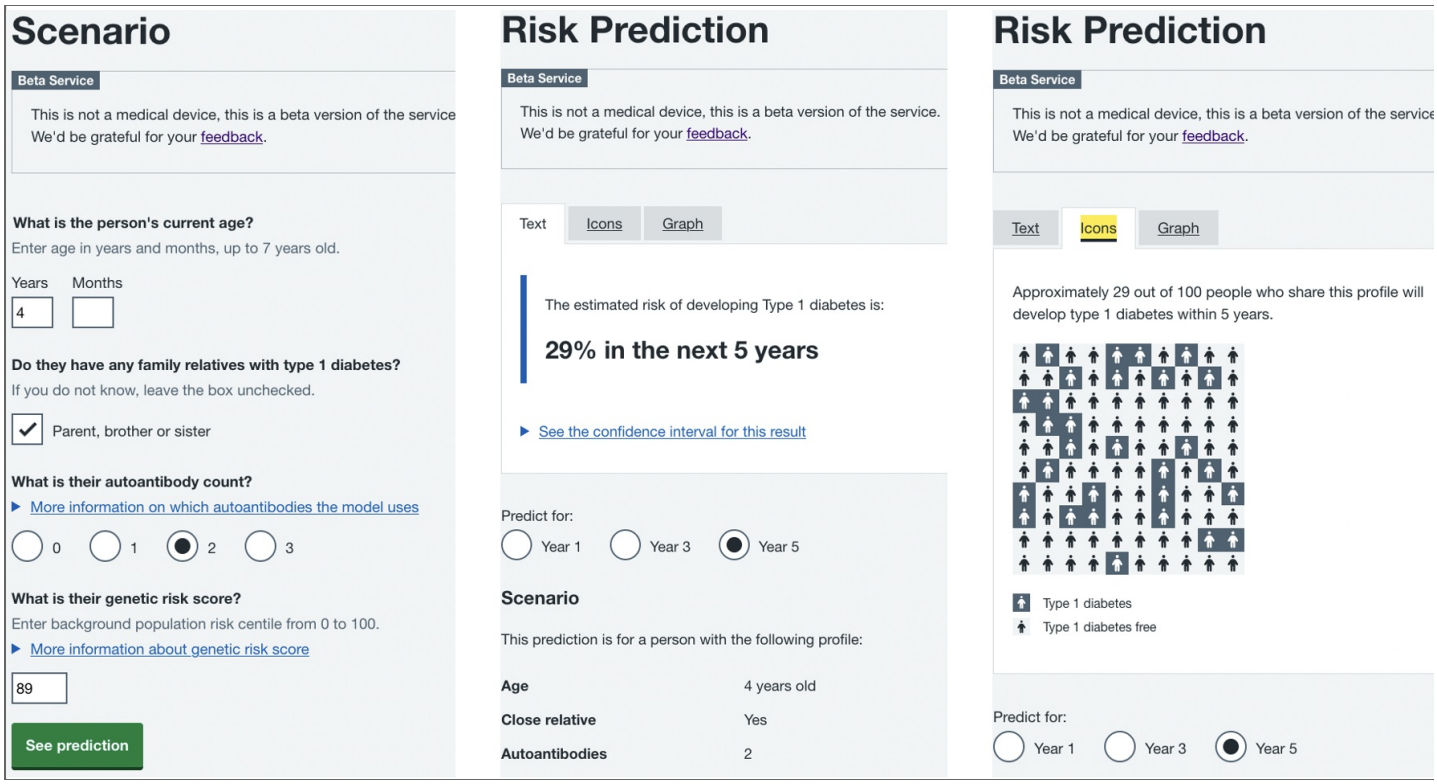
Model patient time at single AB



Stage 1 and 2


Single AB
Stage 1
Stage 2
Brier score
AUC
0.5
0.6
0.7
0.8
0.9
Variables included in the model
Base (clinical + AB)
Base + GRS
Base + Metabolic
Base + GRS + Metabolic
Classic models

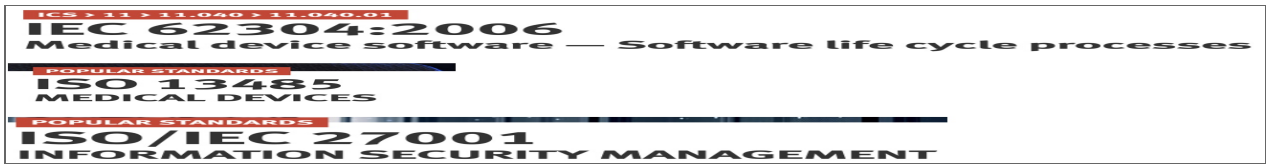
A good model is an accessible model



Model available at **https://t1dpredictor.diabetesgenes.org/**

A good model is an accessible model

Easy to do a prototype
Hard to be legally compliant FDA, GDPR



Conclusion

- Despite its cost and its burden for patient OGTT related measures appear as critical for a good T1D prediction
- Cost and patient time can change drastically for similar predictive performance
- A good model is an accessible model