Received: 26 August 2018 | Revised: 11 November 2018 | Accepted: 4 January 2019

DOI: 10.1111/pedi.12812

WILEY ISPAD

ORIGINAL ARTICLE

Predicting progression to type 1 diabetes from ages 3 to 6 in islet autoantibody positive TEDDY children

Laura M. Jacobsen[1] | Helena E. Larsson[2] | Roy N. Tamura[3] | Kendra Vehik[3] | Joanna Clasen[3] | Jay Sosenko[4] | William A. Hagopian[5] | Jin-Xiong She[6] | Andrea K. Steck[7] | Marian Rewers[7] | Olli Simell[8] | Jorma Toppari[8,9] | Riitta Veijola[10] | Anette G. Ziegler[11] | Jeffrey P. Krischer[3] | Beena Akolkar[12] | Michael J. Haller[1] | the TEDDY Study Group

[1]Department of Pediatrics, University of Florida, Gainesville, Florida

[2]Department of Clinical Sciences Malmö, Lund University, Skåne University Hospital SUS, Malmö, Sweden

[3]Health Informatics Institute, Morsani College of Medicine, University of South Florida, Tampa, Florida

[4]Division of Endocrinology, University of Miami, Miami, Florida

[5]Pacific Northwest Diabetes Research Institute, Seattle, Washington

[6]Center for Biotechnology and Genomic Medicine, Medical College of Georgia, Augusta University, Augusta, Georgia

[7]Barbara Davis Center for Childhood Diabetes, University of Colorado, Denver, Colorado

[8]Department of Pediatrics, Turku University Hospital, Turku, Finland

[9]Department of Physiology, Institute of Biomedicine, University of Turku, Turku, Finland

[10]Department of Pediatrics, Medical Research Center, PEDEGO Research Unit, Oulu University Hospital and University of Oulu, Oulu, Finland

[11]Institute of Diabetes Research, Helmholtz Zentrum München and Forschergruppe Diabetes e.V. Neuherberg, Neuherberg, Germany

[12]Division of Diabetes, Endocrinology, and Metabolism, National Institute of Diabetes, Digestive, and Kidney Diseases, National Institutes of Health, Bethesda, Maryland

Correspondence
Michael J. Haller, MD, P.O. Box 100296, Gainesville, FL 32610.
Email: hallemj@peds.ufl.edu

Funding information
NIH/NCATS Clinical and Translational Science Award University of Colorado, Grant/Award Number: UL1 TR001082; NIH/NCATS Clinical and Translational Science Award University of Florida, Grant/Award Number: UL1 TR000064; TEDDY Study group funding grants, Grant/Award Number: U01 DK63863, U01 DK63836, U01 DK63790 UC4 DK106955, UC4 DK112243, UC4 DK117483 Contract No. HHSN267200700014CU01 DK63829, U01 DK63861, U01 DK63821, U01 DK63865UC4 DK106955, UC4 DK112243, UC4 DK117483UC4 DK63829, UC4 DK63861, UC4 DK63821, UC4 DK63865UC4 DK63863, UC4 DK95300, UC4 DK100238; University of Florida; Florida State University; Columbia University; University of Virginia;

Objective: The capacity to precisely predict progression to type 1 diabetes (T1D) in young children over a short time span is an unmet need. We sought to develop a risk algorithm to predict progression in children with high-risk human leukocyte antigen (HLA) genes followed in The Environmental Determinants of Diabetes in the Young (TEDDY) study.

Methods: Logistic regression and 4-fold cross-validation examined 38 candidate predictors of risk from clinical, immunologic, metabolic, and genetic data. TEDDY subjects with at least one persistent, confirmed autoantibody at age 3 were analyzed with progression to T1D by age 6 serving as the primary endpoint. The logistic regression prediction model was compared to two non-statistical predictors, multiple autoantibody status, and presence of insulinoma-associated-2 autoantibodies (IA-2A).

Results: A total of 363 subjects had at least one autoantibody at age 3. Twenty-one percent of subjects developed T1D by age 6. Logistic regression modeling identified 5 significant predictors - IA-2A status, hemoglobin A1c, body mass index Z-score, single-nucleotide polymorphism rs12708716_G, and a combination marker of autoantibody number plus fasting insulin level. The logistic model yielded a receiver operating characteristic area under the curve (AUC) of 0.80, higher than the two other predictors; however, the differences in AUC, sensitivity, and specificity were small across models.

ABBREVIATIONS: HbA1c, hemoglobin A1c; HLA, human leukocyte antigen; OGTT, oral glucose tolerance test; SNP, single nucleotide polymorphism.

University of Bristol; Davis Center; University of South Florida; **Clinical Center**; Lunds Universitet; **Clinical Center**; TU Dresden; Technische Universität München; **Clinical Center**; University of Florida; Augusta University; **Clinical Center**; Tampere University Hospital; Turku University Hospital; University of Tampere; **Clinical Center**; Davis Center; University of Colorado; **Clinical Center**; University of Colorado; University of Florida; Centers for Disease Control and Prevention; National Institute of Environmental Health Sciences; National Institute of Child Health and Human Development; National Institute of Allergy and Infectious Diseases; National Institute of Diabetes and Digestive and Kidney Diseases

Conclusions: This study highlights the application of precision medicine techniques to predict progression to diabetes over a 3-year window in TEDDY subjects. This multifaceted model provides preliminary improvement in prediction over simpler prediction tools. Additional tools are needed to maximize the predictive value of these approaches.

KEYWORDS

autoantibodies, metabolic, pediatric, prediction, type 1 diabetes

1 | INTRODUCTION

As the worldwide incidence of type 1 diabetes (T1D) continues to increase, there is a growing urgency to develop and test therapies aimed at slowing and stopping progression to T1D in autoantibody positive subjects. While there are currently no effective therapies to stop the development of type 1 diabetes, subgroups within some prevention studies have shown delayed progression and additional promising therapies are currently being explored. In order to most rationally apply these therapies to at-risk children, there is a desire to develop tools to precisely predict disease progression over narrow timeframes.[1,2]

As we move further into the era of precision medicine, the capacity to identify those who are most at-risk for disease development and, perhaps more importantly, those who will respond best to specific therapies continues to progress. That said, efforts to accurately predict and prevent progression to T1D have been underway for some time. The diabetes prevention trial-type 1 (DPT-1) represents the largest single effort to prevent T1D to date.[3] Using the DPT-1 data, the DPT-1 risk score (DPTRS) was developed and validated using body mass index (BMI), age, log-fasting c-peptide, and 2-hour oral glucose tolerance test (OGTT) data to predict T1D risk in DPT-1 participants.[4]

In an effort to build on the concepts of the DPTRS and similar risk score methodologies, we used multivariable logistic regression to assess candidate predictors of T1D progression among young children with one or more positive autoantibodies. The environmental determinants of diabetes in the young (TEDDY) study is a multi-site, multi-country cooperative study aimed at determining which environmental factors are involved in the pathogenesis of T1D.[5] The TEDDY cohort represents a unique group of children with an increased genetic risk of progression to T1D followed since birth. The cohort is of younger age than earlier studies and has unique baseline and longitudinal data that could provide specific information on clinical risk prediction in young children. Such data would be highly valuable in advising high-risk families of the likelihood of disease progression and allowing for early diagnosis. Further benefits of improved risk prediction include avoiding severe complications such as diabetic ketoacidosis (DKA) at diagnosis, which is markedly higher in younger children,[6,7] and aiding in designing and enrolling subjects into prevention trials.

2 | RESEARCH DESIGN AND METHODS

2.1 | Participants and selection criteria

TEDDY subjects were recruited from four countries, the United States, Germany, Sweden, and Finland, and 8676 participants with high-risk human leukocyte antigen (HLA) genotypes were enrolled.[8] Children in the TEDDY study are followed every 3 months until age 48 months for the development of islet autoantibodies, a precursor to T1D development, then every 6 months. For those with autoantibody seroconversion, visits remain every 3 months. Those with 1 autoantibody have a hemoglobin A1c (HbA1c) measured; those with ≥2 autoantibodies additionally have an OGTT starting at age 3 years (2 times points only at 0 and 120 minutes obtained). The requirement for HbA1c measurement with 1 autoantibody was added to the TEDDY study 4 years after the study began. Thus, some subjects did not have HbA1c measurement at age 3. For these subjects, if they had an age 3 HbA1c taken outside of the TEDDY study, this measure was included in the analysis.

Subject data at age 3 was used to determine the risk of progression to T1D by the age of 6 years. Collection of metabolic data, such as HbA1c with or without OGTT begins in the TEDDY study at age 3. As the youngest children in TEDDY have reached 7 years of age, the use of T1D status at age 6 was chosen to ensure complete data collection. All subjects were included if they had ≥1 autoantibody on confirmed testing (2 consecutive positive results) by the age of 3 years 5 months (to include those late for their 3-year visit) but not diagnosed with type 1 diabetes. To determine risk of developing T1D based on predictors at a young age (3 years) and over a short time period (by age 6 years), we included both single and multiple autoantibody positive subjects. We included children at age 3 with single autoantibody status (rather than only looking at those with multiple autoantibodies) as development of additional autoantibodies is more likely in a young population and as these young single autoantibody positive children have higher risk for progression than adolescents or adults with single autoantibodies.[9]

2.2 | Analysis and clinical predictors tested

Thirty-eight variables, determined on the basis of previous TEDDY analyses and available literature, were assessed for association with

progression to T1D by age 6 years (Supporting Information, Table S1). These included clinical characteristics, such as gender, country, general population status vs first-degree relative (including relative specificity), weight, and BMI Z-scores at 3 years of age. General laboratory immunologic and metabolic variables included autoantibody number (single vs multiple), type (glutamic acid decarboxylase autoantibodies [GADA], insulin autoantibodies [mIAA], insulinoma-associated-2 autoantibodies [IA-2A], and zinc transporter 8 autoantibodies [ZnT8A], which was added January 2012 to the protocol), and titer, age at confirmatory autoantibody positivity, and HbA1c. OGTT with two time points (fasting and 2 hours) were performed upon confirmation of multiple autoantibodies and the results placed in clinically-relevant categories based on normal (fasting <90 mg/dL, 2 hours < 120 mg/dL), elevated but not abnormal (fasting 90-100 mg/dL, 2 hours 120-140 mg/dL), or impaired/abnormal (fasting >100 mg/dL, 2 hours > 140 mg/dL) results for blood glucose. Fasting levels for C-peptide, insulin, and homeostatic model assessment (HOMA) to quantify insulin resistance were included. Because of lack of blood sample, levels for C-peptide, insulin, and HOMA were not available at the 2-hour time point for the large majority of subjects. Assessment of fasting and 2-hour glucose values as a continuous variable was performed for a subgroup of subjects at age 3 who had multiple autoantibodies; however, the entire group was analyzed through categorization (Table S1**)** to include those subjects with one autoantibody who did not have an OGTT glucose result. Single nucleotide polymorphisms (SNPs) identified through genome-wide association studies and multiple Cox regression analysis and found to associate with progression to autoantibody positivity and T1D in previous TEDDY analyses were assessed.[10] The SNPs used in this analysis were: rs1004446_A, rs10517086_A, rs11711054_G, rs12708716_G, rs2292239_A, rs2476601_A, rs2816316_C, rs3184504_A, rs3825932_A, rs4948088_A, rs7111341_A, with gene names included in Table S1.

The 38 potential predictor variables were screened for inclusion in the model using a forward selection method with alpha, or significance level, set at 0.01 to enter the model. The level of 0.01 was chosen because of the known increase in false-positive rate when the number of candidate predictor variables becomes large.[11] The operating characteristics, or ability of the model to predict true positives and true negatives, were compared to two simpler models: (a) Does IA-2A status at age 3 predict T1D at age 6, and (b) Does multiple autoantibody status at age 3 predict T1D at age 6? The operating characteristics of the logistic regression model and these two models were assessed by a 4-fold cross-validation procedure.[12] In the cross-validation, the cohort was split into four subsets (folds) of equal size and equal number of T1D cases. Each fold was used as the validation dataset while the remaining subjects were identified as the training dataset. The stepwise logistic regression procedure was generated four times, once for each training dataset and the operating characteristics of the resultant models were then determined on the validation datasets. The final estimates of receiver operator characteristic (ROC) area under the curve (AUC), sensitivity, specificity, positive, and negative predictive values were determined for each validation dataset and summarized by weighted averages over the four validation datasets where the weight was based on the number of subjects in the validation dataset, which had complete data on the chosen predictors. Sensitivity, specificity, positive, and negative predictive values for the logistic regression models were based on the Youden index that maximizes the sum of the sensitivity and specificity. All analyses were performed in PC SAS version 9.3. This study was performed with the approval of the central institutional review board overseeing the TEDDY study.

3 | RESULTS

Of the 8676 participants initially enrolled in TEDDY, 78 developed T1D prior to age 3 years. Another 2235 dropped out of the study by the age of 3 years. Of the remaining subjects at age 3 years, 6000 were autoantibody negative and 363 were autoantibody positive. The subsequent analyses were completed only on the 363 subjects who were autoantibody positive at age 3 years (Figure 1). Of the 363 TEDDY participants who were autoantibody positive at age 3 with complete data, 76 were confirmed to have T1D at age 6. Demographics of both the full cohort and the subset of subjects who progressed are shown in Table 1. The status of 11 subjects in the 363 antibody positive cohort could not be determined at age 6 and these subjects were excluded from analysis. The incidence of T1D by age 6 years was 21.5% (76/352). Plasma glucose >200 mg/dL (>11.1 mmol/L) accounted for 91% of the diagnosed T1D subjects.

Within the TEDDY cohort, logistic regression modeling identified five predictors at age 3 years as significant markers for progression from autoantibody positivity to T1D by age 6 years: presence of IA-2A, HbA1c, BMI Z-score, SNP rs12708716_G, and number of antibodies (multiple autoantibodies) combined with low fasting insulin level (Table 2). The largest effect in the model was seen with IA-2A status at

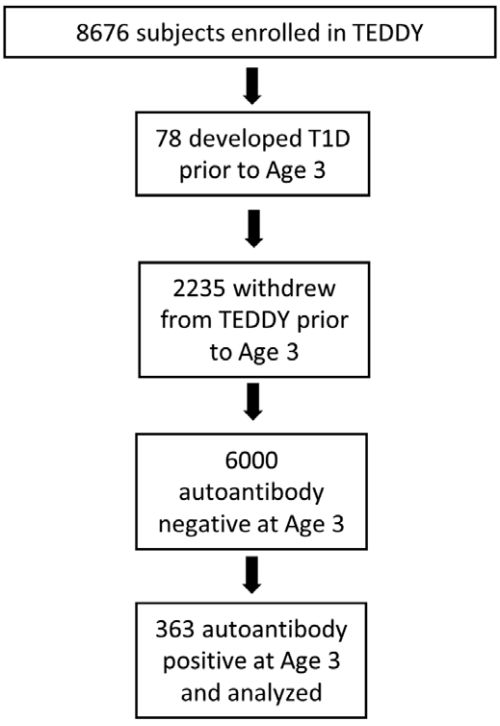

FIGURE 1 Flowchart of children enrolled in cohort for analysis

TABLE 1 Demographics and characteristics at age 3 of complete cohort and the outcome group with type 1 diabetes at age 6 years

	Complete cohort (N = 363) N (%)	T1D subjects: Age 6 (N = 76) N (%)
Age 6 T1D		
Yes	76 (21)	76
No	276 (76)	0
Missing	11 (3)	0
Country		
United States	124 (34)	20 (26)
Finland	83 (23)	24 (32)
Germany	25 (7)	6 (8)
Sweden	131(36)	26 (34)
Gender		
Female	163 (45)	40 (53)
Male	200 (55)	36 (47)
First degree relative		
No	290 (80)	62 (82)
Yes	73 (20)	14 (18)
HLA		
DR3/DR4	177 (49)	43 (57)
DR4/DR4	72 (20)	12 (16)
DR4/DR8	56 (15)	10 (13)
DR3/DR3	43 (12)	7 (9)
Other	15 (4)	4 (5)
Number of autoantibodies		
1	168 (46)	8 (11)
>1	195 (54)	68 (89)
IA-2A status		
Positive	134 (37)	57 (75)
Negative	229 (63)	19 (25)
SNP rs12708716_G number of minor alleles		
0	167 (46)	31 (41)
1	162 (45)	35 (46)
2	32 (9)	11 (13)
Age in years at persistent autoantibody confirmation	**Complete cohort (N = 363)**	**T1D subjects: age 6 (N = 76)**
Mean	1.86	1.53
SD	0.87	0.72
Range	0.25-3.5	0.36-3.3
HbA1c (%)	**Complete cohort (N = 292)**	**All T1D subjects - Age 6 (N = 58)**
Mean	5.14	5.31
SD	0.28	0.27
Range	4.4-6.0	4.7-5.8
BMI (Z-score)	**Complete cohort (N = 329)**	**All T1D subjects – Age 6 (N = 55)**
Mean	0.26	0.42
SD	0.99	1.06
Range	−3.35-2.33	−2.45-1.78

Abbreviations: BMI, body mass index; HLA, human leukocyte antigen; HbA1c, hemoglobin A1c; IA-2A, insulinoma-associated-2 autoantibodies; T1D, type 1 diabetes.

TABLE 2 Logistic regression significant predictors for type 1 diabetes development by age 6

Parameter	Estimate (SE)	*P*-value	Odds ratio (95% CI)
Intercept	−22.7 (4.3)	<0.001	—
Age 3 IA-2A status Reference = negative	2.2 (0.5)	<0.001	8.7 (3.0, 25.2)
Age 3 HbA1c	3.6 (0.8)	<0.001	1.4 (1.2, 1.6)[a]
Age 3 BMI Z Score	0.7 (0.2)	0.002	1.8 (1.2, 2.8)[b]
Number of autoantibodies and OGTT Fasting insulin level Reference = 1 autoantibody	—	0.001	—
>1 antibody—insulin normal	0.6 (0.3)	<0.001	11.7 (2.0, 64.3)
>1 antibody—insulin low	1.5 (0.4)	<0.001	28.5 (4.4, 182.9)
>1 antibody—insulin missing	−0.2 (0.4)	0.605	5.3 (0.8, 32.6)
RS12708716_G	0.9 (0.3)	0.008	2.4 (1.3, 4.5)[c]

Abbreviations: BMI, body mass index; CI, confidence interval; HbA1c, hemoglobin A1c; IA-2A, insulinoma-associated-2 autoantibodies; OGTT, oral glucose tolerance test.
[a] Odds ratio for 0.1 unit increase in HbA1c.
[b] Odds ratio for 1 unit increase in Z-score.
[c] Odds ratio for 1 unit increase in number of alleles (0,1,2).

age 3 (odds ratio [OR] 8.7). An exploratory recursive partition algorithm chose cut points of 5.2% (33 mmol/mol) for HbA1c as optimal in differentiating between high- and low-risk groups for the diagnosis of T1D by age 6. Another metabolic factor that contributed to risk of T1D by age 6 years was a reduced fasting insulin level (<2.0 mcU/mL) in the setting of multiple autoantibodies. In addition, odds for developing T1D by age 6 increased by 1.8 for every unit increase in BMI Z-score. Finally, SNP rs12708716_G (*CLEC16A*), found previously to be a susceptibility locus for islet autoantibody development,[10] demonstrated an OR of 2.4. The Youden index was 0.16 indicating that subjects with an estimated probability of T1D greater than 0.16 are predicted to have T1D by age 6.

Using logistic regression, the two time point OGTT glucose levels were analyzed as continuous variables for only the subset of subjects with 2 or more autoantibodies (n = 114). The 2-hour glucose (OR 1.03 [1.02, 1.05]) but not the fasting glucose (OR 0.98 [0.94, 1.02]) was found to be significantly associated with T1D by age 6.

The estimated ROC AUC, sensitivity, specificity, positive, and negative predictive values for the logistic regression algorithm and the two simpler models are shown in Table 3. In the cross validation, IA-2A status was chosen as the most important predictor in all four of the folds, followed by HbA1c in three of the four folds.

4 | DISCUSSION

Improved prediction of a child's risk of progression to T1D by age 6 has the potential to reduce poor outcomes related to DKA, which occur disproportionately in children under 5 years of age.[7] In addition, time limited predictive tools are urgently needed to effectively deliver personalized medicine and provide families with actionable information when making decisions related to potential participation in prevention studies. Herein, we demonstrated that data collected from 3-year-old children participating in the TEDDY study can be used to

TABLE 3 Operating characteristic of prediction models based on 4-fold cross-validation

Model	ROC AUC	Sensitivity	Specificity	Positive-predictive value	Negative-predictive value
Logistic regression	0.80	0.91	0.59	0.35	0.96
IA-2A yes/no	0.78	0.83	0.73	0.44	0.95
1/>1 positive autoantibody	0.75	0.95	0.56	0.35	0.98

Abbreviations: AUC, area under the curve; IA-2A, insulinoma-associated-2 autoantibodies; ROC, receiver operating characteristic curve.

develop a prediction model for progression to T1D by age 6. Logistic regression modeling provided for the highest estimated AUC; however, the differences in AUC, sensitivity, and specificity between (1) logistic regression modeling, (2) IA-2A status, and (3) single vs multiple autoantibody status, were small. This may be because of the small sample size of the validation datasets, the presence of missing values for HbA1c, and/or the lack of OGTT information for the single autoantibody positive subjects. We compared our combination of factors, derived from 38 candidate predictors to two simple prediction rules—multiple autoantibody status alone or IA-2A positive status alone—to determine if there was added benefit from a model that includes not only autoantibody number, type and titer but also the addition of metabolic data (HbA1c, glucose, insulin), HLA and non-HLA genetic data, gender, country, type of first degree relative, weight, BMI, and more. Logistic regression was chosen as opposed to more algorithmic machine learning techniques because of the relatively small sample size, the relatively low rate of T1D at age 6 and the desire for an interpretable model comparable to other, simpler models.[13] It is difficult to determine whether the drop out of autoantibody negative subjects prior to age 3 is biasing the results of our analysis. With the exception of the rs12708716_G SNP, all of the other predictors selected were based on information collected at age 3, which were unavailable for the subjects who dropped out. While this is a proof-of-concept technique, because of missing data, especially metabolic data at age 3 in single autoantibody subjects, widespread use of this algorithm is not yet justified as simpler models show similarity in operating characteristics. However, additional information can be gained from analysis of the other predictors in addition to autoantibody status. This study highlights the ability to apply precision medicine techniques over a short time period in young children at increased risk of type 1 diabetes.

Birth cohort studies such as the Colorado Diabetes Autoimmunity Study in the Young (DAISY), Germany's BABYDIAB and BABYDIET, and Finland's Type 1 Diabetes Prediction and Prevention (DIPP) have previously demonstrated that progression to T1D is based on age and number of autoantibodies.[14–16] Young children with 2 or more antibodies have a very high lifetime risk of progression to T1D (70% at 10 years and 84% at 15 years).[1,2] Non-birth cohorts, such as the DPT-1, have also confirmed increased risk in autoantibody positive populations.[4,17,18] Similar analyses completed within TEDDY looked at participants with multiple autoantibodies and time to T1D. Age at multiple autoantibody appearance, female sex, and non-HLA SNPs were found to be significant risk factors for time to disease.[19] Our analysis, unlike others, sought to predict progression over a very narrow scope of time in very young children. While the data set from 3 to 6 years of age was limiting, we sought to determine if data collected at age 3 could be used to accurately predict progression over a narrow 3-year time period, which would increase the efficacy of clinical trials. This is an important contrast between our study and previous efforts performed on the entire TEDDY cohort.[19]

Some of the risk predictors identified in our model corroborate earlier studies while others are novel. Steck et al, in previous analyses of TEDDY data, showed that elevated IAA and IA-2A titers were significant risk predictors in young children.[2] In addition, IA-2A has previously been identified as the autoantibody conferring the highest risk of progression. This risk increases with autoantibody titer and epitope reactivity.[2,9,20,21] While IA-2A and multiple autoantibody status were highly correlated, the strongest predictor, IA-2A status, was selected into the model. This confirmed data from previous cohort studies[22] suggesting that IA-2A positivity was a strong predictor of progression in the presence of another autoantibody. These observations are highly pertinent to young at-risk children. IAA status was not significant in the model despite the young age and known early appearance of IAA. Not surprisingly, GADA was not a predictive factor in our analysis for progression to T1D as in DAISY,[23] typically occurring in older subjects.

Excess BMI over time in children within the TrialNet Pathway to Prevention study has demonstrated increased risk of progression that was age and sex-specific.[24] Here, increasing BMI Z score, even in very young children, increased the odds of T1D. HbA1c was an important predictor in our model and has been evaluated previously as a tool for prediction within DAISY,[25] TEDDY,[26] and DIPP.[27] Increase in HbA1c has been shown to be superior to measures of random glucose in predicting progression to T1D.[27] In exploratory analyses, cut-point analysis identified HbA1c at age 3 above 5.2% (33 mmol/mol) in establishing high- and low-risk cohorts with regards to progression at age 6 years. OGTT data based on baseline (fasting) insulin levels in conjunction with single/multiple autoantibody status were a factor in the model. In subgroup analysis of the multiple autoantibody subjects, the 2-hour glucose measurement at age 3 was the most significant OGTT parameter associated with progression by age 6.

The group with single autoantibodies did not have OGTT performed and this is a limitation as artificial categories were made to include the entire cohort studied. When considering the potential application of this, or any risk algorithm, we must always consider the potential logistical concerns of performing the tests required to obtain relevant data in the population of interest. In addition, the TEDDY OGTT provides only a two time point collection (fasting and 2-hour) unlike the DPT-1 or TrialNet Pathway to Prevention protocols which employ a six time point OGTT and is further limited by the fact that 3 years of age is the earliest an OGTT is performed in TEDDY. As such, the relatively limited OGTT dataset in young TEDDY subjects may inhibit our ability to detect signals associated with glucose excursions. Similarly, other variables, such as C-peptide and insulin, for

example, were typically only available in the fasting state without a stimulated value.

Finally, only one SNP previously identified with risk for autoantibody positivity was confirmed as a significant predictor of progression to T1D in our model of very young children. This demonstrates the potential of SNPs to add to risk differentiation well before symptoms or metabolic derangements are detected.[10] The development of autoantibodies early in childhood is associated with HLA class II genotypes whereas those who develop autoimmunity later are less tightly linked to HLA status but non-HLA SNPs may add additional specificity even in young children. Providers caring for young children with a family history of T1D frequently consider autoantibody testing given their well established predictive power. However, future predictive models including type of autoantibody, rising BMI, or rising HbA1c might provide incremental improvements in prediction.

In summary, this report details the performance of logistic regression modeling to predict progression to T1D from age 3 to age 6 in TEDDY children. Given the excellent sensitivity but limited specificity of this model, data from TEDDY children at age 3 has a limited ability to predict progression to T1D by age 6. Explanations for the low specificity and positive predictive value of the prediction model include frequently missing covariates at age 3 (ie, HbA1c), as well as the relatively short interval (3 years) that was utilized in developing this model. Nevertheless, this model provides important proof-of-concept for developing risk scores in very young high-risk children. Improved models are urgently needed in order to stratify children for prevention studies and realize the promise of precision medicine.

ACKNOWLEDGEMENTS

The TEDDY Study Group is funded by U01 DK63829, U01 DK63861, U01 DK63821, U01 DK63865, U01 DK63863, U01 DK63836, U01 DK63790, UC4 DK63829, UC4 DK63861, UC4 DK63821, UC4 DK63865, UC4 DK63863, UC4 DK95300, UC4 DK100238, UC4 DK106955, UC4 DK112243, UC4 DK117483, and Contract No. HHSN267200700014C from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Institute of Child Health and Human Development (NICHD), National Institute of Environmental Health Sciences (NIEHS), Centers for Disease Control and Prevention (CDC), and JDRF. This work supported in part by the NIH/NCATS Clinical and Translational Science Awards to the University of Florida (UL1 TR000064) and the University of Colorado (UL1 TR001082). Dr. Xiang Liu assisted in the classification tree prediction using report. **The TEDDY Study Group: Colorado Clinical Center:** Marian Rewers, M.D., Ph.D., PI[1,4,5,6,10,11], Kimberly Bautista[12], Judith Baxter[9,10,12,15], Ruth Bedoy[2], Daniel Felipe-Morales, Kimberly Driscoll, Ph.D.[9], Brigitte I. Frohnert, M.D.[2,14], Marisa Gallant, M.D.[13], Patricia Gesualdo[2,6,12,14,15], Michelle Hoffman[12,13,14], Rachel Karban[12], Edwin Liu, M.D.[13], Jill Norris, Ph.D.[2,3,12], Adela Samper-Imaz, Andrea Steck, M.D.[3,14], Kathleen Waugh[6,7,12,15], Hali Wright[12]. University of Colorado, Anschutz Medical Campus, Barbara Davis Center for Childhood Diabetes. **Finland Clinical Center**: Jorma Toppari, M.D., Ph.D., PI[¥,1,4,11,14], Olli G. Simell, M.D., Ph.D., Annika Adamsson, Ph.D.[12], Suvi Ahonen[*,±,§], Heikki Hyöty, M.D., Ph.D.[*,±,6], Jorma. Ilonen, M.D., Ph.D.[¥,¶,3], Sanna Jokipuu[^], Tiina Kallio[^], Leena Karlsson[^], Miia Kähönen[µ,¤], Mikael Knip, M.D., Ph.D.[*,±,5], Lea Kovanen[*,±,§], Mirva Koreasalo[*,±,§,2], Kalle Kurppa, M.D., Ph.D.[*,±,13], Tiina Latva-aho[µ,¤], Maria Lönnrot, M.D., Ph.D.[*,±,6], Elina Mäntymäki[^], Katja. Multasuo[µ,¤] Tiina Niininen[±,*,12], Sari Niinistö[±,§,2], Mia Nyblom[*,±], Petra Rajala[^], Jenna. Rautanen[±,§], Anne Riikonen[*,±,§], Minna Romo[^], Juulia Rönkä[µ,¤], Satu Simell, M.D., Ph.D.[¥,13],Tuula Simell, Ph.D.[¥,12], Maija Sjöberg[¥,^,12,14], Aino Stenius[µ,12], Maria Leppänen[^], Sini. Vainionpää[^], Eeva Varjonen[¥,^,12], Riitta Veijola, M.D., Ph.D.[µ,¤,14], Suvi M. Virtanen, M.D., Ph.D.[*,±,§,2], Mari Vähä-Mäkilä[^], Mari Åkerlund[*,±,§], Katri Lindfors, Ph.D.[*,13,¥]University of Turku, [*]University of Tampere, [µ]University of Oulu, [^]Turku University Hospital, Hospital. District of Southwest Finland, [±]Tampere University Hospital, [¤]Oulu University Hospital, [§]National Institute for Health and Welfare, Finland, [¶]University of Kuopio. **Georgia/Florida Clinical Center:** Jin-Xiong She, Ph.D., PI[1,3,4,11], Desmond Schatz, M.D.[*4,5,7,8], Diane Hopkins[12], Leigh Steed[12,13,14,15], Jennifer Bryant, Jamie Thomas[*6,12], Janey Adams[*12], Katherine Silvis[2], Michael Haller, M.D.[*14], Melissa Gardiner, Richard McIndoe, Ph.D., Ashok Sharma, Stephen W. Anderson, M.D.[^], Laura Jacobsen, M.D.[*14]. Center for Biotechnology and Genomic Medicine, Augusta University. [*]University of Florida. [^]Pediatric Endocrine Associates, Atlanta. **Germany Clinical Center:** Anette G. Ziegler, M.D., PI[1,3,4,11], Andreas Beyerlein, Ph.D.[2], Ezio Bonifacio Ph.D.[*,5], Anja Heublein, Michael Hummel, M.D.[13], Sandra Hummel, Ph.D.[2], Annette Knopff[7], Charlotte Koch, Sibylle Koletzko, M.D.[¶13], Claudia Ramminger, Roswith. Roth, Ph.D.[9], Marlon Scholz, Laura Schulzik[2], Joanna Stock[9,12,14], Katharina Warncke, M.D.[14], Lorena Wendel, Christiane Winkler, Ph.D.[2,12,15]. Forschergruppe Diabetes e.V. and Institute of Diabetes Research, Helmholtz Zentrum München, Forschergruppe Diabetes, and Klinikum rechts der Isar, Technische Universität München. [*]Center for Regenerative Therapies, TU Dresden, [¶]Dr. von Hauner Children's Hospital, Department of Gastroenterology, Ludwig Maximillians University Munich. **Sweden Clinical Center:** Åke Lernmark, Ph.D., PI[1,3,4,5,6,8,10,11,15], Daniel Agardh, M.D., Ph.D.[13], Carin Andrén Aronsson, Ph.D.[2,12,13], Maria Ask, Jenny Bremer, Ulla-Marie Carlsson, Corrado Cilio, Ph.D., M.D.[5], Emelie Ericson-Hallström, Annika Fors, Lina Fransson, Thomas. Gard, Rasmus Bennet, Carina Hansson, Susanne Hyberg, Hanna Jisser, Fredrik Johansen, Berglind Jonsdottir, M.D., Silvija Jovic, Helena Elding Larsson, M.D., Ph.D. [6,14], Marielle. Lindström, Markus Lundgren, M.D.[14], Maria Månsson-Martinez, Maria Markan, Jessica. Melin[12], Zeliha Mestan, Caroline Nilsson, Karin Ottosson, Kobra Rahmati, Anita Ramelius, Falastin Salami, Sara Sibthorpe, Anette Sjöberg, Birgitta Sjöberg, Evelyn Tekum Amboh, Carina Törn, Ph.D.[3,15], Anne Wallin, Åsa Wimar[14], Sofie Åberg. Lunds Universitet. **Washington Clinical Center:** William A. Hagopian, M.D., Ph.D., PI[1,3,4,5,6,7,11,13,14], Michael Killian[6,7,12,13], Claire Cowen Crouch[12,14,15], Jennifer Skidmore[2], Ashley Akramoff, Jana Banjanin, Masumeh Chavoshi, Kayleen Dunson, Rachel Hervey, Shana Levenson, Rachel. Lyons, Arlene Meyer, Denise Mulenga, Davey Schmitt, Julie Schwabe. Pacific Northwest. Research Institute. **Pennsylvania Satellite Center:** Dorothy Becker, M.D., Margaret Franciscus, MaryEllen. Dalmagro-Elias Smith[2], Ashi Daftary, M.D., Mary Beth Klein, Chrystal Yates. Children's. Hospital of Pittsburgh of UPMC. **Data Coordinating Center:** Jeffrey P. Krischer, Ph.D.,PI[1,4,5,10,11], Sarah Austin-Gonzalez, Maryouri Avendano, Sandra Baethke, Rasheedah

Brown[12,15], Brant Burkhardt, Ph.D.[5,6], Martha Butterworth[2], Joanna Clasen, David Cuthbertson, Christopher Eberhard, Steven. Fiske[9], Dena Garcia, Jennifer Garmeson, Veena Gowda, Kathleen Heyman, Belinda Hsiao, Francisco Perez Laras, Hye-Seung Lee, Ph.D.[1,2,13,15], Shu Liu, Xiang Liu, Ph.D.[2,3,9,14], Kristian Lynch, Ph.D.[5,6,9,15], Colleen Maguire, Jamie Malloy, Cristina McCarthy[12,15], Aubrie Merrell, Steven Meulemans, Hemang Parikh, Ph.D.[3], Ryan Quigley, Cassandra Remedios, Chris. Shaffer, Laura Smith, Ph.D.[9,12], Susan Smith[12,15], Noah Sulman, Ph.D., Roy Tamura, Ph.D.[1,2,13], Ulla Uusitalo, Ph.D.[2,15], Kendra Vehik, Ph.D.[4,5,6,14,15], Ponni Vijayakandipan, Keith Wood, Jimin Yang, Ph.D., R.D.[2,15]. Past staff: Michael Abbondondolo, Lori Ballard, David. Hadley, Ph.D., Wendy McLeod. University of South Florida. **Project scientist:** Beena Akolkar, Ph.D.[1,3,4,5,6,7,10,11]. National Institutes of Diabetes and Digestive and Kidney Diseases. **Autoantibody Reference Laboratories:** Liping Yu, M.D.[5], Dongmei Miao, M.D.[^], Polly Bingley, M.D., FRCP[*5], Alistair Williams[*], Kyla Chandler[*], Claire Williams[*], Gifty George[*], Sian Grace[*], Ben Gillard[*]. [^]Barbara Davis Center for Childhood Diabetes, University of Colorado Denver, [*]Bristol Medical School, University of Bristol UK. **HLA Reference Laboratory:** William Hagopian[3], M.D., Ph.D., Masumeh Chavoshi, Pacific Northwest Research Institute, Seatte WA. (Previously Henry Erlich, Ph.D.[3], Steven J. Mack, Ph.D., Anna Lisa Fear. Center for Genetics, Children's Hospital Oakland Research Institute.) **SNP Laboratory:** Stephen S. Rich, Ph.D.[3], Wei-Min Chen, Ph.D.[3], Suna Onengut-Gumuscu, Ph. D.[3], Emily Farber, Rebecca Roche Pickin, Ph.D., John Davis, Jordan Davis, Dan Gallo, Jessica Bonnie, Paul Campolieto. Center for Public Health Genomics, University of Virginia. **Repository:** Sandra Ke, Niveen Mulholland, Ph.D. NIDDK Biosample Repository at Fisher BioServices. **Other contributors:** Kasia Bourcier, Ph.D.[5], National Institutes of Allergy and Infectious Diseases. Thomas Briese, Ph.D.[6,15], Columbia University. Suzanne Bennett Johnson, Ph.D.[9,12], Florida State University. Eric Triplett, Ph.D.[6], University of Florida.

Committees:

[1]Ancillary Studies, [2]Diet, [3]Genetics, [4]Human Subjects/Publicity/-Publications, [5]Immune Markers, [6]Infectious Agents, [7]Laboratory Implementation, [8]Maternal Studies, [9]Psychosocial, [10]Quality Assurance, [11]Steering, [12]Study Coordinators, [13]Celiac Disease, [14]Clinical Implementation, [15]Quality Assurance Subcommittee on Data Quality.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to report.

Author contributions

MJH and HEL conceptualized this study and evaluated the data and reviewed/edited the manuscript, LMJ researched the data and wrote the manuscript, RNT and KV analyzed data, contributed to discussion, and reviewed and edited the manuscript, AKS, JS, and WH contributed to discussion and reviewed and edited the manuscript, JXS, JC, MR, OM, JT, RV, AZ, JK, and BA reviewed/edited the manuscript. MJH is the guarantor of this work and takes responsibility for the integrity of the data and the accuracy of the data analysis.

ORCID

Laura M. Jacobsen https://orcid.org/0000-0002-5144-7836
Helena E. Larsson https://orcid.org/0000-0003-3306-1742

REFERENCES

1. Ziegler AG, Rewers M, Simell O, et al. Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. *Jama*. 2013 Jun 19;309(23):2473-2479.
2. Steck AK, Vehik K, Bonifacio E, et al. TEDDY study group. Predictors of progression from the appearance of islet autoantibodies to early childhood diabetes: the environmental determinants of diabetes in the young (TEDDY). *Diabetes Care*. 2015 May;38(5):808-813.
3. Orban T, Sosenko JM, Cuthbertson D, et al. Diabetes prevention trial-type 1 study group. Pancreatic islet autoantibodies as predictors of type 1 diabetes in the diabetes prevention trial-type 1. *Diabetes Care*. 2009 Dec;32(12):2269-2274.
4. Sosenko JM, Krischer JP, Palmer JP, et al. The diabetes prevention trial- type 1 study group. A risk score for type 1 diabetes derived from autoantibody-positive participants in the diabetes prevention trial-type 1. *Diabetes Care*. 2008;31:528-533.
5. TEDDY Study Group. The environmental determinants of diabetes in the young (TEDDY) study. *Ann N Y Acad Sci*. 2008;1150:1-13.
6. Elding Larsson H, Vehik K, Bell R, et al. Reduced prevalence of diabetic ketoacidosis at diagnosis of type 1 diabetes in young children participating in longitudinal follow-up. *Diabetes Care*. 2011;34:2347-2352.
7. Wolfsdorf J, Glaser N, Sperling MA. Diabetic ketoacidosis in infants, children, and adolescents—a consensus statement from the American Diabetes Association. *Diabetes Care*. 2006;29(5):1150-1159.
8. Bonifacio E. Predicting type 1 diabetes using biomarkers. *Diabetes Care*. 2015 Jun;38(6):989-996.
9. Törn C, Hadley D, Lee HS, et al. Role of type 1 diabetes associated SNPs on risk of autoantibody positivity in the TEDDY study. *Diabetes*. 2015 May;64(5):1818-1829.
10. Hui S, Li L. Positive false discover rate estimate in step-wise variable selection. *Commun Stat Simul Comput*. 2007;36:1217-1231.
11. Hatie T, Tibshirani R, Freedman J. *The Elements of Statistical Learning – Data Mining, Inference and Prediction, Section 7.3. Cross-Validiation*. New York, NY: Springer-Verlag; 2001:214-217.
12. Harrell, F. (2018) *Road Map for Choosing between Statistical Modeling and Machine Learning* [Blog post]. Powered by the Academic theme for Hugo, http://www.fharrell.com/post/stat-ml/
13. Barker JM, Barriga KJ, Yu L, et al. Diabetes autoimmunity study in the young. Prediction of autoantibody positivity and progression to type 1 diabetes: diabetes autoimmunity study in the young (DAISY). *J Clin Endocrinol Metab*. 2004 Aug;89(8):3896-3902.
14. Kimpimäki T, Kulmala P, Savola K, et al. Natural history of beta-cell autoimmunity in young children with increased genetic susceptibility to type 1 diabetes recruited from the general population. *J Clin Endocrinol Metab*. 2002 Oct;87(10):4572-4579.
15. Hummel S, Ziegler AG. Early determinants of type 1 diabetes: experience from the BABYDIAB and BABYDIET studies. *Am J Clin Nutr*. 2011 Dec;94(suppl_6):1821S-1823S.
16. Sosenko JM, Skyler JS, Mahon J, et al. Type 1 diabetes TrialNet and diabetes prevention trial-type 1 study groups. Use of the diabetes prevention trial-type 1 risk score (DPTRS) for improving the accuracy of the risk classification of type 1 diabetes. *Diabetes Care*. 2014 Apr;37 (4):979-984.
17. Sosenko JM, Skyler JS, Mahon J, et al. Validation of the diabetes prevention trial–type 1 risk score in the TrialNet natural history study. *Diabetes Care*. 2011 Aug;34(8):1785-1787.
18. Krischer J, Liu X, Lernmark A, et al. The influence of type 1 diabetes genetic susceptibility regions, age, sex, and family history to the progression from multiple autoantibodies to type 1 diabetes: a TEDDY study report. *Diabetes*. 2017 Dec;66(12):3122-3129.
19. Ziegler AG, Nepom GT. Prediction and pathogenesis in type 1 diabetes. *Immunity*. 2010 Apr 23;32(4):468-478.
20. Michels A, Zhang L, Khadra A, Kushner JA, Redondo MJ, Pietropaolo M. Prediction and prevention of type 1 diabetes: update

on success of prediction and struggles at prevention. *Pediatr Diabetes*. 2015 Nov;16(7):465-484.
21. Decochez K, De Leeuw IH, Keymeulen B, et al. IA-2 autoantibodies predict impending type I diabetes in siblings of patients. *Diabetologia*. 2002 Dec;45(12):1658-1666.
22. Steck AK, Johnson K, Barriga KJ, et al. Age of islet autoantibody appearance and mean levels of insulin, but not GAD or IA-2 autoantibodies, predict age of diagnosis of type 1 diabetes: diabetes autoimmunity study in the young. *Diabetes Care*. 2011;34(6):1397-1399.
23. Ferrara CT, Geyer SM, Liu YF, et al. Type 1 diabetes TrialNet study group. Excess BMI in childhood: a modifiable risk factor for type 1 diabetes development? *Diabetes Care*. 2017 May;40(5):698-701. https://doi.org/10.2337/dc16-2331 Epub 2017 Feb 15.
24. Stene LC, Barriga K, Hoffman M, et al. Normal but increasing hemoglobin A1c levels predict progression from islet autoimmunity to overt type 1 diabetes: diabetes autoimmunity study in the young (DAISY). *Pediatr Diabetes*. 2006 Oct;7(5):247-253.
25. Vehik K, Cuthbertson D, Boulware D, et al. Performance of HbA1c as an early diagnostic indicator of type 1 diabetes in children and youth. *Diabetes Care*. 2012;35:1821-1825.
26. Helminen O, Aspholm S, Pokka T, et al. HbA1c predicts time to diagnosis of type 1 diabetes in children at risk. *Diabetes*. 2015;64:1719-1727.
27. Veijola R, Koskinen M, Helminen O, Hekkala A. Dysregulation of glucose metabolism in preclinical type 1 diabetes. *Pediatr Diabetes*. 2016; 17 (Suppl. 22), 25-30.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Jacobsen LM, Larsson HE, Tamura RN, et al. Predicting progression to type 1 diabetes from ages 3 to 6 in islet autoantibody positive TEDDY children. *Pediatr Diabetes*. 2019;1–8. https://doi.org/10.1111/pedi.12812

Supplemental Table 1: List of 38 candidate predictors used in analyses.

Variable
HLA Category, 5 Categories
DR3/DR4
DR4/DR4
DR4/DR3
DR3/DR3
Other
Gender
Country
First Degree Relative - Any
First Degree Relative - Mother
First Degree Relative - Father
First Degree Relative - Sibling
Age 3 GADA Autoantibody Titer Z score
Age 3 IA2A Autoantibody Titer Z score
Age 3 mIAA Autoantibody Titer Z score
Age 3 ZnT8A Autoantibody Titer Z score
Age 3 GADA Positive (yes/no)
Age 3 IA2A Positive (yes/no)
Age 3 mIAA Positive (yes/no)
Age 3 ZnT8A Positive (yes/no)
Age 3 Single vs. Multiple Positive Antibodies
Age 3 HbA1c
Age 3 Change from prior HbA1c
Age 3 Weight
Age 3 BMI
Age at First Persistent Confirmed Autoantibody Positive
Probiotics less than 28 days (yes/no)
Age 3 Number of Antibodies / OGTT Glucose (mg/dL), 5 Categories:
1 Antibody+
>1 Antibody+ Normal Glucose (fasting < 90, 2hr < 120)
>1 Antibody+ Elevated Glucose (fasting 90-100, 2hr 120-140)
>1 Antibody+ Impaired Glucose (fasting > 100, 2hr > 140)
>1 Antibody+ Unknown Glucose (OGTT missing)
Age 3 Number of Antibodies / OGTT Total Glucose (mg/dL), 5 Categories:
1 Antibody+
>1 Antibody+ Fasting + 2 hr Glucose ≤ 200
>1 Antibody+ Fasting + 2 hr Glucose > 200 and ≤ 230
>1 Antibody+ Fasting + 2 hr Glucose > 230
>1 Antibody+ Unknown Glucose (OGTT missing)
Age 3 Number of Antibodies / OGTT Fasting C-peptide, 4 Categories:
1 Antibody+
>1 Antibody+ Normal C-peptide (fasting ≥ 0.6 ng/mL)
>1 Antibody+ Low C-peptide (fasting < 0.6 ng/mL)
>1 Antibody+ Unknown C-peptide (OGTT missing)
Age 3 Number of Antibodies / OGTT Fasting Insulin, 4 Categories:
1 Antibody+
>1 Antibody+ Normal Insulin (fasting ≥ 2.0 mcU/mL)
>1 Antibody+ Low Insulin (fasting < 2.0 mcU/mL)
>1 Antibody+ Unknown insulin (OGTT missing)
Age 3 Number of Antibodies / OGTT Fasting HOMA, 4 Categories:
1 Antibody+

>1 Antibody+ Low HOMA (fasting ≤ 1.0)
>1 Antibody+ High HOMA (fasting > 1.0)
>1 Antibody+ Unknown HOMA (OGTT missing)
rs1004446_A – Gene: *INS*
rs10517086_A – Gene: Unknown
rs11711054_G – Gene: *CCR5*
rs12708716_G – Gene: *CLEC16A*
rs2292239_A – Gene: *ERBB3*
rs2476601_A – Gene: *PTPN22*
rs2816316_C – Gene: *RGS1*
rs3184504_A – Gene: *SH2B3*
rs3825932_A – Gene: *CTSH*
rs4948088_A – Gene: *COBL*
rs7111341_A – Gene: *INS*