



Islet Autoimmunity and HLA Markers of Presymptomatic and Clinical Type 1 Diabetes: Joint Analyses of Prospective Cohort Studies in Finland, Germany, Sweden, and the U.S.

Diabetes Care 2021;44:1–8 | <https://doi.org/10.2337/dc20-1836>

Vibha Anand,¹ Ying Li,¹ Bin Liu,¹
 Mohamed Ghalwash,^{1,2} Eileen Koski,¹
 Kenney Ng,¹ Jessica L. Dunne,³
 Josefina Jönsson,⁴ Christiane Winkler,^{5,6,7}
 Mikael Knip,^{8,9,10,11} Jorma Toppari,¹²
 Jorma Ilonen,¹³ Michael B. Killian,¹⁴
 Brigitte I. Frohnert,¹⁵ Markus Lundgren,⁵
 Anette-Gabriele Ziegler,^{5,6,7}
 William Hagopian,¹⁴ Riitta Veijola,¹⁶ and
 Marian Rewers,¹⁵
 for the T1DI Study Group*

OBJECTIVE

To combine prospective cohort studies, by including HLA harmonization, and estimate risk of islet autoimmunity and progression to clinical diabetes.

RESEARCH DESIGN AND METHODS

For prospective cohorts in Finland, Germany, Sweden, and the U.S., 24,662 children at increased genetic risk for development of islet autoantibodies and type 1 diabetes have been followed. Following harmonization, the outcomes were analyzed in 16,709 infants-toddlers enrolled by age 2.5 years.

RESULTS

In the infant-toddler cohort, 1,413 (8.5%) developed at least one autoantibody confirmed at two or more consecutive visits (seroconversion), 865 (5%) developed multiple autoantibodies, and 655 (4%) progressed to diabetes. The 15-year cumulative incidence of diabetes varied in children with one, two, or three autoantibodies at seroconversion: 45% (95% CI 40–52), 85% (78–90), and 92% (85–97), respectively. Among those with a single autoantibody, status 2 years after seroconversion predicted diabetes risk: 12% (10–25) if reverting to autoantibody negative, 30% (20–40) if retaining a single autoantibody, and 82% (80–95) if developing multiple autoantibodies. HLA-DR-DQ affected the risk of confirmed seroconversion and progression to diabetes in children with stable single-autoantibody status. Their 15-year diabetes incidence for higher- versus lower-risk genotypes was 40% (28–50) vs. 12% (5–38). The rate of progression to diabetes was inversely related to age at development of multiple autoantibodies, ranging from 20% per year to 6% per year in children developing multipositivity in ≤ 2 years or > 7.4 years, respectively.

CONCLUSIONS

The number of islet autoantibodies at seroconversion reliably predicts 15-year type 1 diabetes risk. In children retaining a single autoantibody, HLA-DR-DQ genotypes can further refine risk of progression.

¹Center for Computational Health, IBM T.J. Watson Research Center, Cambridge, MA, and Yorktown Heights, NY

²Ain Shams University, Cairo, Egypt

³JDRF, New York, NY

⁴Department of Clinical Sciences Malmö, Lund University/CRC, Skåne University Hospital, Malmö

⁵Institute of Diabetes Research, Helmholtz Zentrum München German Research Center for Environmental Health, Munich-Neuherberg, Germany

⁶Forschergruppe Diabetes e.V. am Helmholtz Zentrum München, Munich, Germany

⁷Forschergruppe Diabetes, Technical University Munich, Germany

⁸Tampere Center for Child Health Research, Tampere University Hospital, Tampere, Finland

⁹Pediatric Research Center, Children's Hospital, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

¹⁰Research Program for Clinical and Molecular Metabolism, Faculty of Medicine, University of Helsinki, Helsinki, Finland

¹¹Folkhälsan Research Center, Helsinki, Finland

¹²Institute of Biomedicine and Population Research Centre, University of Turku, and Department of Pediatrics, Turku University Hospital, Turku, Finland

¹³Immunogenetics Laboratory, Institute of Biomedicine, University of Turku, and Clinical Microbiology, Turku University Hospital, Turku, Finland

¹⁴Pacific Northwest Research Institute, Seattle, WA

¹⁵Barbara Davis Center for Diabetes, University of Colorado, Denver, CO

¹⁶PEDEGO Research Unit, Department of Pediatrics, University of Oulu and Oulu University Hospital, Oulu, Finland

Corresponding author: Vibha Anand, anand@us.ibm.com

Received 21 July 2020 and accepted 7 April 2021

This article contains supplementary material online at <https://doi.org/10.2337/figshare.14384324>.

Type 1 diabetes is a chronic autoimmune endocrine disease that affects an estimated 1 in 300 children and up to 1 in 120 adults in the U.S. (1) and more in high-risk Nordic countries (2,3). The causes of the underlying islet autoimmunity are poorly understood, and no durable prevention or cure is available. The genetic and environmental determinants of type 1 diabetes have been extensively investigated in cohort studies that followed children at increased genetic risk for development of islet autoimmunity and progression to diabetes (4–8). Observations from these cohorts (9) led to the staging of the natural history of type 1 diabetes: stage 1, normoglycemia with presence of multiple islet autoantibodies; stage 2, dysglycemia; and stage 3, clinical (symptomatic) diabetes (10).

While the average annual rate of progression from stage 1 to stage 3 is ~11% (11), the individual risk is difficult to predict due to large variability in the progression rate (5,8,12–14). To overcome this limitation, investigators for the five active cohort studies in Finland (Diabetes Prediction and Prevention Study [DIPP]) (6), Germany (BABYDIAB) (5), Sweden (Diabetes Prediction in Skåne Study [DiPiS]) (8), and the U.S. (Diabetes Autoimmunity Study in the Young [DAISY] [4] and Diabetes Evaluation in Washington [DEW-IT] [7]) harmonized and combined their data for joint analyses in collaboration with IBM Research and JDRF—known as the Type 1 Diabetes Intelligence (T1DI) Study Group. We are reporting the risk estimates for development of islet autoimmunity and progression to clinical diabetes with stratification by the number of autoantibodies and HLA-DR-DQ genotype in children followed from infancy.

Additionally, we are reporting the subsequent risk of progression to clinical diabetes with stratification by age at seroconversion. In contrast to an earlier report (9) with stratification of the risk by the maximum number of antibodies ever expressed over an extended period of time, here the risk is stratified by the number of autoantibodies observed at the time of initial seroconversion. This approach better reflects information available to a screening program.

RESEARCH DESIGN AND METHODS

Study Populations in the Combined Data Set

The prospective cohort studies included in this report enrolled 1) infants from the general population identified through newborn screening as carrying increased-risk HLA-DR-DQ genotypes (4,6–8) and first-degree relatives of people with type 1 diabetes (regardless of HLA genotype [4,5]). All study participants underwent HLA screening; however, the eligibility criteria varied (see Supplementary Material [Study Sites, Newborn Screening and Recruitment]). Age at the initial follow-up visit ranged from 2 months to 21.6 years; participants were followed at 3- to 36-month intervals for development of islet autoantibodies, according to the study-specific protocols, for up to 26 years. Type 1 diabetes was diagnosed according to the American Diabetes Association criteria (15).

All individual study protocols were approved by local institutional review boards, and the sites submitted deidentified data to IBM Research in accordance with the Health Insurance Portability and Accountability Act and General Data Protection Regulation. IBM Research aggregated and harmonized the data and performed the analyses.

T1DI Study Cohorts

Overall Harmonized Cohort

The visits with positive autoantibody measurements deemed to be maternally transferred were excluded. Children lacking evidence of positivity for any of the islet autoantibodies prior to diagnosis of diabetes were excluded ($n = 54$); most were lost to follow-up early, only to be found diagnosed with type 1 diabetes years after their last study visit. The overall harmonized cohort for analyses included 24,662 subjects with 285,217 study visits, with a median of 10 visits and 8.7 years of follow-up per subject. Of those, 4,165 (17%) subjects reported family history of type 1 diabetes. Further characteristics of this cohort are shown in Table 1 and Supplementary Tables 1 and 2.

Infant-Toddler Cohort

For key analyses presented in this report, we selected a more homogeneous subcohort from the overall cohort, referred to hereafter as the infant-toddler cohort. (See Supplementary Fig. 1 for summary of cohort selection.) The infant-toddler cohort only includes subjects who were initially tested for islet autoantibodies to insulin, IA-2, and GAD at or before 2.5 years of age. DIPP participants who were born prior to 2003 were excluded ($n = 4,297$), as they were only screened with islet cell antibodies assay. The infant-toddler cohort includes 16,709 subjects with 215,757 study visits, with a median 12 visits per subject and 10.4 years of follow-up (Table 1).

Data Variables and HLA Harmonization

A minimal set of common features was extracted and standardized from the submitted data sets (Supplementary Table 1). The HLA genotypes were harmonized across these studies (see *HLA Risk Groups* below). Subject-level (static) features included date of birth month and year, sex, family history and relationship with type 1 diabetes proband, HLA-DR-DQ genotype, breastfeeding (ever), and age at diagnosis of type 1 diabetes. The visit-level (dynamic) variables included age at each visit and autoantibodies (titer level and positive/negative outcome) to insulin, GAD, IA-2, and zinc transporter 8 (ZnT8A) as well as height, weight, plasma glucose levels, and HbA_{1c}. The individual titer levels for antibody assays were not harmonized for this study and will be presented in our future work. Instead, the binary outcomes, i.e., (positive/negative) of autoantibody measurement, submitted by each study site, were used for this study. Similarly, data for standardized height (to centimeters), weight (to kilograms), HbA_{1c} (to NGSP, U.S. standard), and glucose levels (to milligrams per deciliter) will be presented in our future work. Sociodemographics, breastfeeding, and family history (and relationship) where known are described in Supplementary Table 2 by individual study site.

*A full list of study sites for the study group can be found below in the Appendix. In addition, see supplementary material online for the full list of study group members.

© 2021 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered.

More information is available at <https://www.diabetesjournals.org/content/license>. See accompanying article, p. XXX.

Table 1—T1DI study cohorts

Study, site, enrollment period, age at enrollment	Age range, [median age] at follow-up visits, years	Cohort type	N cohort subjects (no. of visits)	N with confirmed seroconversion (% of cohort) [†]	N with multiple autoantibodies ever (% of cohort)	N with T1D (% of cohort)	N with T1D excluded [‡]
BABYDIAB, Germany, 1989–2000, 9 months	0.0–28.5 [2.1]	Overall	2,364 (27,179)	220 (9.0)	123 (5.0)	95 (4.0)	12
BABYDIAB, Germany, 1989–2000, 9 months	0.0–28.5 [2.1]	Infant-toddler	2,346 (27,130)	220 (9.0)	123 (5.0)	95 (4.0)	0
DAISY, CO, 1993–2006, 9 months*	0.6–29.8 [7.4]	Overall	2,539 (26,803)	199 (8.0)	129 (5.0)	101 (4.0)	5
DAISY, CO, 1993–2006, 9 months*	0.6–25.1 [7.0]	Infant-toddler	2,170 (23,402)	165 (8.0)	105 (5.0)	81 (3.7)	0
DEW-IT, WA, 1995–2001, 2008–2012	0.4–45.2 [8.1]	Overall	3,748 (9,196)	173 (5.0)	170 (5.0)	56 (1.5)	8
DEW-IT, WA, 1995–2001, 2008–2012	0.4–16.5 [2.2]	Infant-toddler	559 (1,490)	17 (3.0)	18 (3.0)	8 (1.4)	0
DiPIS, Sweden, 2000–2004, 24 months	0.0–13.0 [4.0]	Overall	4,359 (34,298)	184 (4.0)	100 (2.0)	75 (1.7)	1
DiPIS, Sweden, 2000–2004, 24 months	0.0–13.0 [4.0]	Infant-toddler	4,353 (34,280)	184 (4.0)	100 (2.0)	75 (1.7)	0
DIPP, Finland, 1994–2009, 2–6 months [§]	0.0–22.9 [4.9]	Overall	11,652 (187,741)	837 (7.0)	526 (5.0)	399 (3.4)	28
DIPP, Finland, 1994–2009, 2–6 months [§]	0.0–22.9 [5.0]	Infant-toddler	7,281 (129,455)	827 (11.0)	519 (7.0)	396 (5.4)	0
All	0.0–45.2 [4.9]	Overall	24,662 (285,217)	1,613 (7.0)	1,048 (4.2)	726 (2.9)	54
All	0.0–28.5 [4.6]	Infant-toddler	16,709 (215,757)	1,413 (8.4)	865 (5.2)	655 (3.9)	0

*First degree relatives were eligible to enroll if younger than 8 years (born 1993–1995) or younger than 4 years (1996–2004). [†]Seroconversion per T1DI study cohort definition in *Research Design and Methods*, without inclusion of ZnT8A positivity. [‡]Those with T1D without autoantibody measurement or seropositivity in the follow-up period. [§]DIPP subjects born before 2003 were screened with only islet cell antibodies assay (N = 4,297). ^{||}Without ZnT8A in the autoantibody count.

HLA Risk Groups

Genotypes from individual studies were harmonized into four risk groups: A, B, C, D (ordered by decreasing risk, e.g., A = DR4-DQ8/DR3-DQ2.5 represents the highest risk). This harmonization was performed based on prior risk information (16) of HLA-DRB1, -DQA1, and -DQB1 alleles (17–19). Five broad “haplotype groups” were defined as follows: DR3-DQ2 included DQB1*02 positivity together with DQA1*05. All DQB1*03:02-positive subjects were identified as positive for DR4-DQ8, and all other haplotypes were grouped as either neutral (X), protective (Y), or highly protective (Z) (16,20,21). See Supplementary Table 3 for haplotype groups. For each subject, two haplotypes were individually assigned to one of the five haplotype groups and then together mapped to one of the four HLA risk groups as described in Supplementary Table 4. In the infant-toddler cohort, 2,212 (14%) subjects were assigned to HLA risk group A, 6,632 (40%) to group B, 2,508 (15%) to group C, and 5,179 (31%) to group D. A total of 178 subjects (<1%) could not be assigned to any HLA risk group because of missing genotype information (missing at least one haplotype or both and/or a missing allele) and were excluded for analyses. In the infant-toddler cohort, the proportion of subjects with the highest-risk genotypes (HLA group A) was higher in the U. S. cohorts, 22% in DAISY and 28% in DEW-IT, than in the European cohorts: 13% in DiPIS, 12% in DiPIS, and 6% in BABYDIAB. In BABYDIAB, only first-degree relatives were enrolled and HLA genotypes were not used as eligibility criteria, albeit HLA typing was performed. Please see Supplementary Table 5 for details of HLA risk group assignment. Furthermore, note that of the 3,525 (21%) subjects with family history of type 1 diabetes in the infant-toddler cohort, 254 (7%) were assigned to HLA group A, 801 (23%) to group B, 739 (21%) to group C, and 1,577 (45%) to group D and 154 (4%) subjects remained unassigned.

Islet Autoantibodies

Methods used for each study to measure islet autoantibodies to insulin, IA-2, and GAD are summarized in Supplementary Material in a section on measurement of islet autoantibodies. These assays have evolved greatly over the past 26 years. For each of the studies rigorous quality control procedures were used to control for a drift in the assays and their laboratories have participated with satisfactory results in all concurrent Diabetes Autoantibody Standardization Program (DASP) (22) and its successor Islet Autoantibody Standardization

program (IASP) (23) proficiency workshops. For this study, a binary result (positive/negative) was produced for each islet autoantibody measurement.

Study End Points and Islet Autoimmunity Definitions

The primary study end points in the infant-toddler cohort were as follows:

- Confirmed seroconversion, defined as positivity for the same islet autoantibody at two or more consecutive visits regardless of the interval between the visits. Confirmed seroconversion age was defined as the age at the first of the consecutive positive visits.
- Positivity for multiple islet autoantibodies at confirmed seroconversion or subsequently; the age at positivity for multiple islet autoantibodies was defined as the age when the second autoantibody was first detected.
- Clinical diabetes.

Islet autoimmunity was primarily defined by the number of positive autoantibodies (i.e., 1, 2, or 3) at confirmed seroconversion. Separately, for children with a single autoantibody at seroconversion, we assessed autoantibody status 2 years postseroconversion based on a recent report (24) of persistence or reversion. Therefore, we report on development of islet autoimmunity at the analytic time point of 2 years past confirmed seroconversion as follows: S-0, single autoantibody at confirmed seroconversion with reversion to no antibodies 2 years later; S-S, single autoantibody at confirmed seroconversion with no subsequent development of any

additional antibody, referred to as “stable single”; and S-M, single autoantibody at confirmed seroconversion with development of multiple antibodies within 2 years. While for all five cohort studies ZnT8A were measured, these were not included in the analyses, as they were generally measured only if the subject tested positive for one or more of the other three autoantibodies or had developed diabetes. Inclusion of ZnT8A did not change the overall results (data not shown).

Statistical Analyses

Survival analyses were performed to generate cumulative risk estimates by number of autoantibodies at seroconversion (and 2 years later if single autoantibody-positive) and compare those by age and HLA risk groups. Age at development of multiple islet autoantibodies was stratified into quartiles for comparisons of the annual incidence rates of progression to diabetes (11). For all analyses of progression to diabetes, event time was defined as the age at diagnosis of clinical diabetes or the age at last follow-up visit for those who did not progress. Kaplan-Meier curves were plotted with 95% CIs, and the log-rank test was used to test for statistical differences. With stratification by HLA risk groups, or by number of islet autoantibodies, pairwise statistical comparisons were made. We calculated positive predictive value (PPV) and sensitivity (SENS) of number of islet autoantibodies to development of type 1 diabetes by using inverse probability of censoring weighting (25) to handle censored observations. Significance was tested at $P < 0.05$,

and analyses were conducted with Python v3.6 and statistical package R.

RESULTS

In the infant-toddler cohort of 16,709 subjects, 1,413 (8.4%) had confirmed seroconversion and 865 (5.2%) developed multiple autoantibodies (stage 1 type 1 diabetes) (Table 1). The median age at confirmed seroconversion was 4.0 years, and it was 3.8 years for development of multiple autoantibodies in children who developed multiple autoantibodies at seroconversion or thereafter. Overall, 655 (3.9%) children were diagnosed with clinical stage 3 type 1 diabetes.

Incidence of Islet Autoimmunity and Clinical Diabetes by HLA Risk Group

In the infant-toddler cohort, the risk of confirmed seroconversion to a single autoantibody, multiple autoantibodies, and clinical diabetes differed by the HLA risk groups ($P < 0.0001$ for all three comparisons). By the age of 15 years, confirmed seroconversion to a single autoantibody occurred in 14% (95% CI 12.5–17.5) of children in HLA group A (DR3-DQ2.5/DR4-DQ8.1), 9% (8.0–10.0) in group B, 8% (7.0–9.0) in group C, and 7.5% (7.5–11.0) in group D (Fig. 1A). A large proportion (45%) of children in group D were first-degree relatives. Confirmed seroconversion to multiple autoantibodies by age 15 years occurred in 7% (6.0–9.0) of children in HLA risk group A, 2.5% (2.0–3.0) in group B, 1% (0.5–1.5) in group C, and 1% (0.5–2.0) in group D (Fig. 1B).

Clinical diabetes developed by age 15 years in 16% (95% CI 14.0–18.0) of

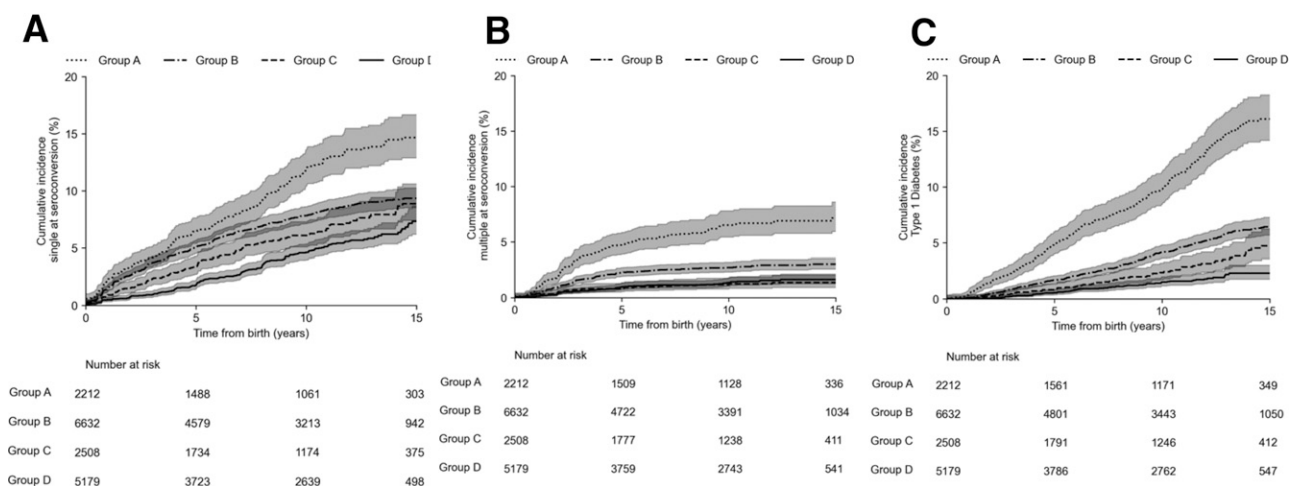


Figure 1—A: Cumulative incidence of confirmed seroconversion to a single islet autoantibody from birth by HLA risk group in the infant-toddler cohort. B: Cumulative incidence of confirmed seroconversion to multiple islet autoantibodies from birth by HLA risk group in the infant-toddler cohort. C: Cumulative incidence of type 1 diabetes from birth by HLA risk group in the infant-toddler cohort.

children in HLA risk group A, 6% (5.0–7.0) in group B, 4% (3.0–5.0) in group C, and in 2.5% (2.0–3.0) in group D (Fig. 1C).

Type 1 Diabetes Incidence by Number of Autoantibodies

Among children with confirmed seroconversion in the infant-toddler cohort (N = 1,413), the majority (N = 1,047 [75%]) were single autoantibody-positive at seroconversion. The 15-year cumulative incidence of diabetes from seroconversion varied (P < 0.0001) in those with one (N = 1,047), two (N = 281), or three (N = 85) autoantibodies at confirmed seroconversion: 45% (95% CI 40–52), 85% (78–90), and 92% (85–97), respectively (Fig. 2A).

The PPV of one, two, or three autoantibodies to develop type 1 diabetes in 15 years after seroconversion was 0.53 (95% CI 0.47–0.59), 0.67 (0.60–0.73), and 0.71 (0.63–0.79), respectively. Similarly, the SENS of one, two, or three autoantibodies to develop type 1 diabetes in 15 years since seroconversion was 0.35 (0.33–0.38), 0.40 (0.38–0.42), and 0.42 (0.39–0.44), respectively.

The 15-year cumulative incidence of diabetes in children seroconverting with a single autoantibody varied significantly (P < 0.0001) when 2 years’ follow-up postseroconversion was considered (Fig. 2B). The 15-year cumulative incidence among those with S-0, S-S, and S-M status was 12% (95% CI 10–25), 30% (20–40), and 82% (80–95), respectively.

The PPV of single autoantibody status at 2 years postseroconversion, i.e., S-0,

S-S, and S-M, to develop type 1 diabetes in 15 years was 0.21 (95% CI 0.11–0.31), 0.31 (0.21–0.41), and 0.67 (0.55–0.79), respectively. Similarly, the SENS of S-0, S-S, and S-M to develop type 1 diabetes in 15 years was 0.18 (0.11–0.25), 0.25 (0.19–0.30), and 0.40 (0.36–0.44), respectively.

Furthermore, among children positive with a stable single autoantibody (S-S) in the infant-toddler cohort, the 15-year cumulative incidence of diabetes varied significantly (P < 0.0001) by higher- versus lower-risk HLA genotypes: 40% (95% CI 28.0–50.0) for HLA group A or B vs. 12% (5.0–38.0) for HLA group C or D genotypes (Fig. 2C). Please also see Supplementary Fig. 4B for stratification by individual HLA risk groups A–D in the infant-toddler cohort.

A summary of the cumulative incidence of diabetes in the overall cohort with stratification by HLA risk groups and the number of islet autoantibodies at seroconversion (and 2 years postseroconversion) is given in Supplementary Table 6.

Type 1 Diabetes Incidence by Age at Development of Multiple Islet Autoantibodies

Of the 865 subjects who developed multiple autoantibodies in the infant-toddler cohort, 812 had confirmed seroconversion (at least 2 consecutive positive visits). For analysis of the effect of age on the rate of progression to clinical diabetes, age at development of multiple autoantibodies was categorized in quartiles: 1Q, ≤2.0 years (n = 202, of

whom 177 have developed diabetes); 2Q, 2.0–3.8 years (n = 205, 151 with diabetes); 3Q, 3.8–7.4 years (n = 202, 126 with diabetes); and 4Q, 7.4–18 years (n = 203, 59 with diabetes). During the initial 10 years of follow-up since development of multiple autoantibodies, the overall incidence of diabetes was 12 per 100 person-years. This rate decreased (P < 0.0001) with increasing age quartile (1Q–4Q) to 20, 12, 11, and 6 per 100 person-years, respectively (Table 2 and Supplementary Fig. 5). However, the annual incidence rate was stable within each age quartile over time. In the infant-toddler cohort, the 10-year cumulative incidence of diabetes in children with multiple autoantibodies was double in the ≤2 years age-group in comparison with the >7.4 years age-group (87% vs. 44%) (Table 2).

CONCLUSIONS

We combined and harmonized data from five prospective cohorts of type 1 diabetes in the U.S. and Europe in a single data set. From these we generate robust risk estimates for development of Islet autoimmunity and progression to clinical diabetes by number of islet autoantibodies and HLA-DR-DQ genotypes for up to 15 years of follow-up. In contrast to the earlier report (9) that retrospectively defined Islet autoimmunity based on the maximum number of antibodies ever expressed between birth and age 15 years, our analysis is based on the number of autoantibodies

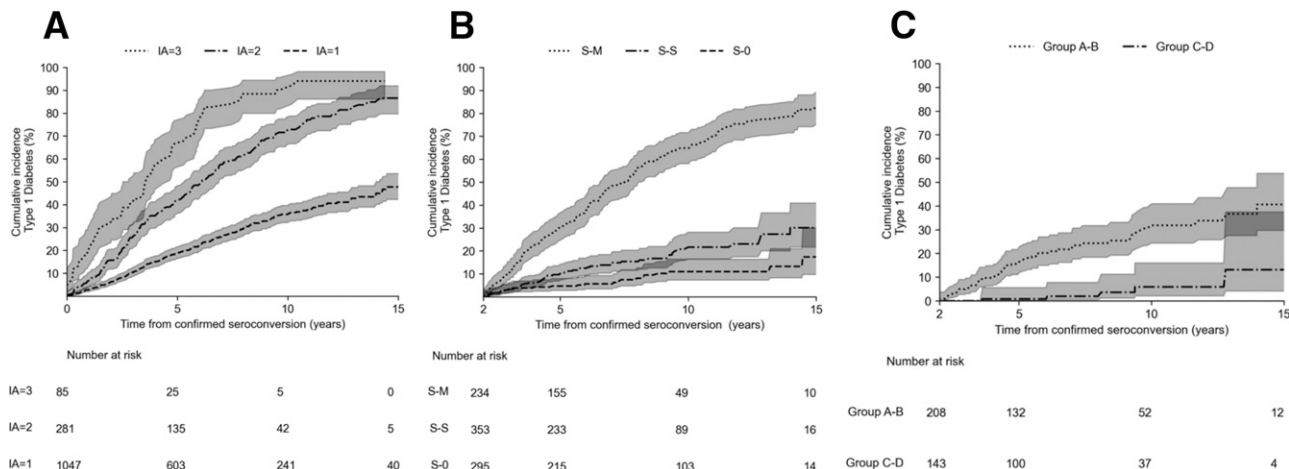


Figure 2—A: Cumulative incidence of type 1 diabetes by the number of positive islet autoantibodies (IA) (one, two, or three) at confirmed seroconversion in the infant-toddler cohort. B: Cumulative incidence of type 1 diabetes among subjects with a single autoantibody at confirmed seroconversion in the infant-toddler cohort. C: Cumulative incidence of type 1 diabetes by HLA risk group among subjects with a stable single autoantibody (S-S) 2 years postseroconversion in the infant-toddler cohort.

Table 2—Risk of progression to type 1 diabetes among children positive for multiple islet autoantibodies in infant-toddler cohort (N = 812)

Quartiles of age distribution at development of multiple islet autoantibodies	N	Cumulative 5-year incidence of diabetes	Cumulative 10-year incidence of diabetes	Cumulative 15-year incidence of diabetes	Average annual incidence of diabetes over 10 years (per 100, per year)
≤2.0 years	202	64% (57–70%)	87% (80–91%)	95% (89–98%)	19.9
>2.0 and ≤3.8 years	205	42% (35–49%)	72% (64–77%)	82% (74–88%)	12.4
>3.8 and ≤7.4 years	202	35% (28–42%)	67% (58–74%)	83% (72–90%)	10.8
>7.4 and ≤18.1 years	203	28% (20–35%)	44% (32–54%)	54% (36–67%)	5.6
Total	812				12.2

observed at the time of initial seroconversion and 2 years later. The baseline for the estimated risk of progression to diabetes in the previous report (9) was the age when study subjects achieved maximum autoantibodies, while in the current report it is the age at initial seroconversion. The latter more closely represents the risk from the perspective of a screening program, where risk prediction cannot include information that is not yet available. The two reports defined risk by number of islet autoantibodies and the baseline for follow-up in a different way and yielded different risk estimates, an important consideration in the context of diabetes risk counseling.

While this previous work also analyzed a combined data set from three of the studies (DIPP, BABYDIAB, and DAISY) that were used in our combined cohort, the analyses presented here go beyond in several important ways. We include two additional cohorts (DiPiS and DEW-IT) and up to 6 more years of follow-up of the study participants from the previous studies. To our knowledge, the T1DI study cohort is the largest data set of prospectively collected information concerning predictors of childhood type 1 diabetes, with 24,662 children followed for up to 26 years and a subcohort of 16,709 children followed since infancy.

Three major findings are shown herein. First, children who initially seroconvert to multiple autoantibodies had greater cumulative risk of type 1 diabetes than those who initially develop a single autoantibody, despite the fact that many in the latter group have subsequently developed multiple autoantibodies. While this is consistent with previous observations with risk stratification by the maximum number of autoantibodies

ever expressed (9,11,26), our findings highlight the importance of the earliest biomarkers in the evolution of Islet autoimmunity. We have estimated PPV and SENS of number of autoantibodies at seroconversion for development of type 1 diabetes in 15 years while accounting for censoring in the past observational studies. These findings highlight the importance of risk evaluation based on single autoantibody versus multiple autoantibodies development at the earliest time point in the course of Islet autoimmunity development.

Second, the younger the age of multiple autoantibodies appearance, the greater the rate of progression to clinical type 1 diabetes—consistent with previous studies (9,27). However, this was found both in the infant-toddler cohort and in the overall cohort (see Supplementary Table 7 for overall cohort).

Third, the HLA-DR-DQ genotype significantly influences progression to diabetes among children seroconverting and remaining positive for a stable single autoantibody at least 2 years past seroconversion. Previous reports (28,29) have missed this effect by not subdividing initially single autoantibody-positive children according to their status 2 years later. However, we confirmed previous reports that the HLA effect was negligible in those with multiple autoantibodies at seroconversion.

A concern in screening for childhood risk of type 1 diabetes is regarding those with a single autoantibody at seroconversion. Their overall rate of developing type 1 diabetes during childhood is substantial (30%) but much lower than the rate among those with multiple autoantibodies (>80%). Our findings suggest that addition of genetic markers and a repeat islet autoantibody test 2 years later may improve individual

risk assessment in the single autoantibody group. Here, we show that HLA-DR-DQ genotypes may be useful in this regard. Interestingly, our analyses also suggest that single autoantibody-positive children in HLA risk group C or D develop diabetes at a later age than those in group A or B. Thus, our findings emphasize that at confirmed seroconversion, the single autoantibody-positive subjects (75% in this large cohort) have a substantially lower rate of progression to diabetes compared with multiple autoantibodies subjects. Among those remaining positive for a single autoantibody, the risk can be stratified based on HLA-DR-DQ genotypes. Among those with multiple autoantibodies, the risk can be stratified by age at development of Islet autoimmunity. We believe these findings can positively inform recruitment in prevention trials and pave way for screening protocols.

Advantages of this report include the large data set representing populations at moderate (Germany and the U.S.) to high (Finland and Sweden) risk of type 1 diabetes and children followed from birth for up to 26 years. Substantial input from multiple investigators representing these studies made it possible to harmonize and jointly analyze the data. Harmonization of the HLA-DRB1-DQA1-DQB1 genotypes across the five cohorts was an unprecedented challenge. In the populations studied, a limited number of stable haplotypes was expected (30) so that even when not all three loci were typed, it was usually possible to infer the specific haplotypes. Then, using disease odds ratios from large collections of cases like the Type 1 Diabetes Genetics Consortium (16), we showed that it is readily possible to assign HLA genotype risk groups. For islet

autoantibody tests, laboratories serving each study have long participated in the DASP (22) and its successor the IASP (23). Consistent participation in the proficiency workshops at 18-month intervals has allowed the laboratories to adhere to standardized quality control procedures and monitor the accuracy of the assays, leading to broadly comparable islet autoantibody data.

Our study has some limitations. The study population is predominantly Caucasian, the harmonized data set does not contain information on non-HLA genotypes though some of the participating studies submitted these data and this may expand. Subjects eligible for prospective follow-up had increased HLA-conferred genetic susceptibility, but subjects were also included on the basis of positive family history of type 1 diabetes. It may more often be that the latter carry non-HLA susceptibility genes in comparison with what has been observed in the population. The majority of our HLA risk group D subjects were positive for family history, and that is probably reflected in their diabetes risk, which was clearly higher than in the general population. Since the HLA genotyping in the original studies was crude (circa 2000), there is also a far broader representation of HLA genotypes in our cohort than can be found in a typical preselected cohort for type 1 diabetes study (31), and in that regard our cohort is a bit more like a general background population. We did not evaluate the type or order of appearance of autoantibodies or specific autoantibody combinations in this study. Understanding relation of various islet autoimmunity profiles to genetic background will be a focus for our future work. The order of appearance of autoantibodies has been shown to be related to HLA-DR-DQ genotype, at least in one study (32). The T1DI Study Group agreed to use binary outcomes of autoantibody titers for the current analyses, and the titer values are being harmonized for forthcoming manuscripts. The original study protocols included somewhat different eligibility criteria and follow-up visit frequency. Longer intervals between study visits hamper identification of the true seroconversion time. The Environmental Determinants of Diabetes in the Young (TEDDY) (31) will overcome this limitation; 8,676 high-risk children recruited at 3–4 months of age will be followed with islet autoantibody assessment every 3 months in the initial 4 years of life and every 6

months thereafter until age 15 years. When the entire TEDDY cohort passes the 15-year mark, in late 2024, the data set will provide higher-resolution answers regarding seroconversion and risk of clinical diabetes. The T1DI study illustrates the variability in approaches to screening and follow-up for childhood diabetes in diverse settings in the U.S. and Europe. However, it also provides a proof of principle that such a diverse data set can be harmonized and jointly analyzed to generate robust risk estimates for children and adolescents. In contrast, very few data on islet autoimmunity and genetic markers are currently available in the adult population.

Our results are generally consistent with the published literature. In the past, it has been difficult to generalize results from specific birth cohort studies due to marked differences in study populations, eligibility criteria, and follow-up protocols. Our large and HLA-harmonized data set is already being used to explore more granular patterns of the development of islet autoantibodies (type, timing, and titer) and dysglycemia in relation to HLA and family history background, sex, growth, geography, and diet. Future application of novel analytical methods (33) such as machine learning and data-driven approaches should increase our understanding of type 1 diabetes pathogenesis and prediction. This may include the application of tools already developed in other settings to visualize data-driven clusters (34) and disease progression models (35,36). These approaches require large and diverse data sets such as those of the T1DI cohort that we hope will pave the way to a more precise approach to prediction and prevention of type 1 diabetes.

Acknowledgments. The authors thank the participants of the DAISY, DiPiS, DIPP, DEW-IT, and BABYDIAB studies. Frank Martin is a representative of JDRF and the convener and funder of the overall initiative.

Funding. This work was supported by funding from JDRF (IBM Research, 1-RSC-2017-368-I-X, 1-IND-2019-717-I-X; DAISY, 1-SRA-2019-722-I-X, 1-RSC-2017-517-I-X, 5-ECR-2017-388-A-N; DiPiS, 1-SRA-2019-720-I-X, 1-RSC-2017-526-I-X; DIPP, 1-RSC-2018-555-I-X; and DEW-IT, 1-SRA-2019-719-I-X, 1-RSC-2017-516-I-X) as well as Center for Scientific Review, National Institutes of Health (DAISY, DK032493, DK032083, DK104351, and DK116073, and DiPiS, DK26190) and the Centers for Disease Control and Prevention (DEW-IT, UR6/CCU017247). The DIPP study was funded by JDRF (grants 1-SRA-2016-342-M-R, 1-SRA-2019-732-M-B), European Union (grant BMH4-CT98-3314), Novo Nordisk

Foundation, Academy of Finland (decision no. 292538 and Centre of Excellence in Molecular Systems Immunology and Physiology Research 2012–2017, decision no. 250114), Special Research Funds for University Hospitals in Finland, Diabetes Research Foundation, Finland, and Sigrid Juselius Foundation, Finland. BABYDIAB was funded by the German Federal Ministry of Education and Research to the German Center for Diabetes Research. DiPiS was funded by Swedish Research Council (grant no. 14064), Swedish Childhood Diabetes Foundation, Swedish Diabetes Association, Nordisk Insulin Fund, The Research Funds of Skane University Hospital, Lion Club International (district 101-S), the Royal Physiographic Society, and Skåne County Council Foundation for Research and Development and received LUDC-IRC/EXODIAB funding from the Swedish Foundation for Strategic Research (Dnr IRC15-0067) and Swedish Research Council (Dnr 2009-1039). Additional funding for DEW-IT was provided by the Hussman Foundation and by the Washington State Life Science Discovery Fund.

Duality of Interest. V.A., M.G., E.K., and K.N. are current employees of IBM. Y.L. and B.L. also are former employees of IBM and performed this work while at IBM. J.L.D. performed this work as an employee of JDRF and is now an employee of Janssen, Inc. No other potential conflicts of interest relevant to this article were reported.

Author Contributions. V.A. is Principal Investigator for the IBM Research study, wrote the first and final draft of the manuscript, and is the technical research lead for this project and takes responsibility for all analyses presented. Y.L. and B.L. helped with subanalyses. M.G. helped with feature engineering (seroconversion and multiple positivity). E.K., K.N., and J.L.D. provided logistical support for the project and draft revisions of the manuscript and were responsible for the initial conceptualization of the T1DI cohort study. J.I. and W.H. helped with the HLA harmonization scheme. M.K., J.T., B.I.F., M.L., A.-G.Z., W.H., M.L., R.V., and M.R. are site primary and coinvestigators and were responsible for collecting study data. J.J., C.W., and M.B.K. helped with site-specific data collection and support processes. All authors made substantial contributions to conception and design of the manuscript, participated in drafting the manuscript or revising it critically for important intellectual content, and gave final approval of the version to be submitted. M.L., A.-G.Z., W.H., R.V., and M.R., representatives of the data-originating sites, are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract form at the 79th Scientific Sessions of the American Diabetes Association, San Francisco, CA, 7–11 June 2019.

APPENDIX

T1DI Study Group study sites: Center for Computational Health at IBM Research and JDRF and DAISY, DiPiS, DIPP, DEW-IT, and BABYDIAB study sites.

References

- Maahs DM, West NA, Lawrence JM, Mayer-Davis EJ. Epidemiology of type 1 diabetes. *Endocrinol Metab Clin North Am* 2010;39:481–497
- Berhan Y, Waernbaum I, Lind T, Möllsten A; Swedish Childhood Diabetes Study Group. Thirty years of prospective nationwide incidence of childhood type 1 diabetes: the accelerating increase by time tends to level off in Sweden. *Diabetes* 2011;60:577–581
- Harjutsalo V, Sund R, Knip M, Groop P-H. Incidence of type 1 diabetes in Finland. *JAMA* 2013;310:427–428
- Rewers M, Bugawan TL, Norris JM, et al. Newborn screening for HLA markers associated with IDDM: diabetes autoimmunity study in the young (DAISY). *Diabetologia* 1996;39:807–812
- Ziegler AG, Hummel M, Schenker M, Bonifacio E. Autoantibody appearance and risk for development of childhood diabetes in offspring of parents with type 1 diabetes: the 2-year analysis of the German BABYDIAB Study. *Diabetes* 1999;48:460–468
- Kupila A, Muona P, Simell T, et al.; Juvenile Diabetes Research Foundation Centre for the Prevention of Type 1 Diabetes in Finland. Feasibility of genetic and immunological prediction of type 1 diabetes in a population-based birth cohort. *Diabetologia* 2001;44:290–297
- Wion E, Brantley M, Stevens J, et al. Population-wide infant screening for HLA-based type 1 diabetes risk via dried blood spots from the public health infrastructure. *Ann N Y Acad Sci* 2003;1005:400–403
- Elding Larsson H. A Swedish approach to the prevention of type 1 diabetes. *Pediatr Diabetes* 2016;17(Suppl. 22):73–77
- Ziegler AG, Rewers M, Simell O, et al. Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. *JAMA* 2013;309:2473–2479
- Insel RA, Dunne JL, Atkinson MA, et al. Staging presymptomatic type 1 diabetes: a scientific statement of JDRF, the Endocrine Society, and the American Diabetes Association. *Diabetes Care* 2015;38:1964–1974
- Bonifacio E. Predicting type 1 diabetes using biomarkers. *Diabetes Care* 2015;38:989–996
- Skyler JS. Characterizing subgroups of type 1 diabetes. *Diabetes* 2014;63:3578–3580
- Arif S, Leete P, Nguyen V, et al. Blood and islet phenotypes indicate immunological heterogeneity in type 1 diabetes. *Diabetes* 2014;63:3835–3845
- 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
- American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2014;37(Suppl. 1):S81–S90
- Erich H, Valdes AM, Noble J, et al.; Type 1 Diabetes Genetics Consortium. HLA DR-DQ haplotypes and genotypes and type 1 diabetes risk: analysis of the Type 1 Diabetes Genetics Consortium families. *Diabetes* 2008;57:1084–1092
- Siljander HTA, Simell S, Hekkala A, et al. Predictive characteristics of diabetes-associated autoantibodies among children with HLA-conferred disease susceptibility in the general population. *Diabetes* 2009;58:2835–2842
- Wenzlau JM, Liu Y, Yu L, et al. A common nonsynonymous single nucleotide polymorphism in the SLC30A8 gene determines ZnT8 autoantibody specificity in type 1 diabetes. *Diabetes* 2008;57:2693–2697
- Salonen KM, Ryhänen S, Härkönen T, Ilonen J; Finnish Pediatric Diabetes Register. Autoantibodies against zinc transporter 8 are related to age, metabolic state and HLA DR genotype in children with newly diagnosed type 1 diabetes. *Diabetes Metab Res Rev* 2013;29:646–654
- Emery LM, Babu S, Bugawan TL, et al. Newborn HLA-DR,DQ genotype screening: age- and ethnicity-specific type 1 diabetes risk estimates. *Pediatr Diabetes* 2005;6:136–144
- Ilonen J, Reijonen H, Herva E, et al. Rapid HLA-DQB1 genotyping for four alleles in the assessment of risk for IDDM in the Finnish population. The Childhood Diabetes in Finland (DiMe) Study Group. *Diabetes Care* 1996;19:795–800
- Bingley PJ, Bonifacio E, Mueller PW. Diabetes Antibody Standardization Program: first assay proficiency evaluation. *Diabetes* 2003;52:1128–1136
- Lampasona V, Pittman DL, Williams AJ, et al.; Participating Laboratories. Islet Autoantibody Standardization Program 2018 Workshop: *Inte. Clin Chem* 2019;65:1141–1152
- Vehik K, Lynch KF, Schatz DA, et al.; TEDDY Study Group. Reversion of β -cell autoimmunity changes risk of type 1 diabetes: TEDDY study. *Diabetes Care* 2016;39:1535–1542
- Vock DM, Wolfson J, Bandyopadhyay S, et al. Adapting machine learning techniques to censored time-to-event health record data: a general-purpose approach using inverse probability of censoring weighting. *J Biomed Inform* 2016;61:119–131
- 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;38:808–813
- Krischer JP, Liu X, Lernmark Å, et al.; TEDDY Study Group. The influence of type 1 diabetes genetic susceptibility regions, age, sex, and family history on the progression from multiple autoantibodies to type 1 diabetes: a TEDDY study report. *Diabetes* 2017;66:3122–3129
- Bingley PJ; European Nicotinamide Diabetes Intervention Trial (ENDIT) Group. Progression to type 1 diabetes in islet cell antibody-positive relatives in the European Nicotinamide Diabetes Intervention Trial: the role of additional immune, genetic and metabolic markers of risk. *Diabetologia* 2006;49:881–890
- Koskinen MK, Lempainen J, Löytyniemi E, et al. Class II HLA genotype association with first-phase insulin response is explained by islet autoantibodies. *J Clin Endocrinol Metab* 2018;103:2870–2878
- Klitz W, Maiers M, Spellman S, et al. New HLA haplotype frequency reference standards: high-resolution and large sample typing of HLA DR-DQ haplotypes in a sample of European Americans. *Tissue Antigens* 2003;62:296–307
- Hagopian WA, Erlich H, Lernmark A, et al.; TEDDY Study Group. The Environmental Determinants of Diabetes in the Young (TEDDY): genetic criteria and international diabetes risk screening of 421 000 infants. *Pediatr Diabetes* 2011;12:733–743
- Krischer JP, Lynch KF, Schatz DA, et al.; TEDDY Study Group. The 6 year incidence of diabetes-associated autoantibodies in genetically at-risk children: the TEDDY study. *Diabetologia* 2015;58:980–987
- Dugas M. Clinical research informatics: recent advances and future directions. *Yearb Med Inform* 2015;10:174–177
- Kwon BC, Eysenbach B, Verma J, et al. Clustervision: visual supervision of unsupervised clustering. *IEEE Trans Vis Comput Graph* 2018;24:142–151
- Kwon BC, Achenbach P, Dunne JL, et al. Modeling disease progression trajectories from longitudinal observational data. *AMIA Annu Symp Proc* 2020;668-676
- Kwon BC, Anand V, Severson KA, et al. DPVis: visual analytics with hidden Markov models for disease progression pathways. *IEEE Trans Vis Comput Graph*. 7 April 2020 [Epub ahead of print]. DOI: 10.1109/TVCG.2020.2985689