

RESEARCH ARTICLE

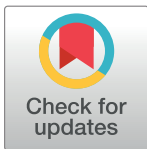
# Sibling method increases risk assessment estimates for type 1 diabetes

Hoang V. Lam<sup>1</sup>, Dat T. Nguyen<sup>2</sup>, Cao D. Nguyen<sup>3,4\*</sup>

**1** Department of Endocrinology, Cho Ray Hospital, Ho Chi Minh City, Vietnam, **2** Department of Science and Technology, Hoa Sen University, Ho Chi Minh City, Vietnam, **3** Department of Business and Information System, Economics University, Ho Chi Minh City, Vietnam, **4** Clinical Analysis and Modelling—Department of Health, Western Australia, Australia

☞ These authors contributed equally to this work.

\* [cao.nguyen@health.wa.gov.au](mailto:cao.nguyen@health.wa.gov.au)



## Abstract

We presented a risk assessment model to distinguish between type 1 diabetes (T1D) affected and unaffected siblings using only three single nucleotide polymorphism (SNP) genotypes. In addition we calculated the heritability from genome-wide identity-by-descent (IBD) sharing between full siblings. We analyzed 1,253 pairs of affected individuals and their unaffected siblings (750 pairs from a discovery set and 503 pairs from a validation set) from the T1D Genetics Consortium (T1DGC), applying a logistic regression to analyze the area under the receiver operator characteristic (ROC) curve (AUC). To calculate the heritability of T1D we used the Haseman-Elston regression analysis of the squared difference between the phenotypes of the pairs of siblings on the estimate of their genome-wide IBD proportion. The model with only 3 SNPs achieving an AUC of 0.75 in both datasets outperformed the model using the presence of the high-risk DR3/4 HLA genotype, namely AUC of 0.60. The heritability on the liability scale of T1D was approximately from 0.53 to 0.92, close to the results obtained from twin studies, ranging from 0.4 to 0.88.

## OPEN ACCESS

**Citation:** Lam HV, Nguyen DT, Nguyen CD (2017) Sibling method increases risk assessment estimates for type 1 diabetes. PLoS ONE 12(5): e0176341. <https://doi.org/10.1371/journal.pone.0176341>

**Editor:** Petter Bjornstad, University of Colorado Denver School of Medicine, UNITED STATES

**Received:** December 22, 2016

**Accepted:** April 10, 2017

**Published:** May 16, 2017

**Copyright:** © 2017 Lam et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** The authors received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

One of the main reasons for disease gene identification is to provide the ability to identify people who are at risk of disease. Thus, a central question for the field is whether validated marker data can be used to discriminate effectively between cases and controls. However, even markers with replicated highly significant odds ratios may be poor classifiers and most variants identified so far confer only small increments in risk and still explain only a small proportion of phenotypic diversity [1]. T1D is a major chronic childhood disease caused by a combination of genetic and environmental influences and genome wide association studies (GWAS) have found over 60 genes to affect the risk of the disease, with the HLA loci having the greatest impact on susceptibility (reviewed in [2, 3]). However, the AUC for risk prediction using multiple identified variants ranges from 0.65 to 0.68 for T1D (see ref. [4] for more details) despite the fact that T1D has a very strong family component with a heritability estimate from 0.4 to 0.8 [5–8].

The association of T1D with alleles at HLA loci, especially the HLA class II genes DR and DQ, is well validated [9]. The highest risk is seen in individuals who are heterozygous HLA-DRB1\*03 and HLA-DRB1\*04 types. HLA allele typing assists in determining risk for T1D, and in studies to understand the pathogenesis of T1D. It is particularly useful in prevention and intervention trials that test potential preventative treatments in high-risk subjects [10]. However, the high cost of HLA genotyping is a major impost on such large scale programs but is beyond the reach of smaller research groups.

In this study, we presented a cost-effective predictive model that could distinguish T1D status in siblings from multiplex families. Our model can be conducted at birth for early prediction and prevention. Our 3-SNP model can not only prevent mortality, but also decrease morbidity and public health costs.

## Materials and methods

### T1D datasets

We used subjects from the Type 1 Diabetes Genetics Consortium (T1DGC) [11]. A subject was labelled as affected if the subject had documented T1D with onset at 37 years old, had used insulin within 6 months of diagnosis, and had no concomitant disease or disorder associated with diabetes. Most subjects came from families where more than one child was affected, and genotyping and clinical data were also collected for parents and unaffected sibs.

For each family, we randomly selected an affected subject to form a dataset, namely, *probands*. Next, sibs were selected from each family and were paired together with the probands to create 1,253 pairs of *proband-sib*. Then we randomly split the 1,253 pairs into two datasets, namely, a “discovery” dataset of 750 pairs and a “validation” dataset of 503 pairs, subject to the equal proportion of case vs control in each dataset.

### Predictors

We have recently presented a 3-SNP set, namely, *rs2854275*, *rs3104413* and *rs9273363* that could rapidly define the HLA-DR and HLA-DQ types relevant to T1D (see our methods in [12]). We used these SNPs genotyped from the probands as well as theirs sibs to predict the risk of a new sibling at birth to be developed T1D in a multiplex family.

### Risk assessment model

We used a logistic regression model to construct risk prediction models [13]. This method finds the logistic curve that best predicts the risk of disease  $P = \text{yes}$  on the basis of continuous or categorical independents of an observation  $G = (g_1, \dots, g_n)$ , formally:

$$\text{Ln} \left( \frac{\text{Pr}(P = \text{yes}|G)}{1 - \text{Pr}(P = \text{yes}|G)} \right) = B_0 + B_1g_1 + \dots + B_n g_n \quad (1)$$

Logistic regression uses an approach called maximum likelihood estimation to estimate regression coefficients. In this case, the *predictor*  $G$  is a 9-dimension vector, including the 3 SNPs genotyped from a proband, the corresponding 3 SNPs genotyped from a sib and 3 binary indicators showing whether or not a genotype from the proband is equal to the corresponding genotype from the sib.

We measured the discriminative accuracy of the predictive models using receiver-operator curve (ROC) analyses [14,15]. The ROC plots the relationship between the true positive rate (TPR or sensitivity) and false positive rate (FPR or 1-specificity) across all possible threshold values that define the disease. The area under the receiver-operator curve (AUC) is the

probability that a randomly chosen case will have a higher estimated risk of developing the disease than a randomly chosen control. The AUC ranges from 0.5 to 1, where a higher number implies a better discriminative model between cases and controls. One important feature of the AUC is that it is not dependent on the number of cases or controls tested as described in ref. [16].

### Heritability

Heritability of disease traits is formally defined as the proportion of phenotypic variance in a population attributable to additive genetic factors [17]. Traditionally, heritability is often estimated on the basis of parent-offspring correlations for continuous traits or the ratio of the incidence in first-degree relatives of affected persons to the incidence in first-degree relatives of unaffected persons [5]. In this study, we used the Haseman-Elston regression analysis [18], the simple estimation procedure for  $n$  sib pairs, of the squared difference between the phenotypes  $P_{i1}$  and  $P_{i2}$  of the  $i^{th}$  pair of siblings on the estimate of their genome-wide IBD proportion  $\hat{\pi}$ , formally:

$$(P_{i1} - P_{i2})^2 = \alpha + \beta\hat{\pi} \tag{2}$$

Because T1D subjects were recruited from different families across worldwide [2,3] the equation (2) is adjusted for stratification using a linear mixed model:

$$(P_{i1} - P_{i2})^2 = \alpha + \beta\hat{\pi} + u \tag{3}$$

where  $u$  is random effect from T1D samples's regions. The fixed effect coefficient  $\beta$  is estimated using the *lme4* package [19] after determining that the genome-wide IBD were distributed normally. We assume that the parents are not inbred so an estimate of the narrow sense heritability is simply:

$$\hat{h}_o^2 = -\hat{\beta}/(2\hat{\delta}_p^2) \tag{4}$$

where  $\hat{\delta}_p^2$  is an estimate of the total phenotypic variance.

To account for ascertainment that generates a much higher proportion of cases in our analyzed samples than in the population, the estimate heritability  $\hat{h}_o^2$  on the observed scale case-control study was transformed to that on the liability scale as [17,20]:

$$h_l^2 = \hat{h}_o^2 \frac{K(1-K)}{z^2} \frac{K(1-K)}{P(1-P)} \tag{5}$$

Where  $K$  is the population prevalence of T1D disease,  $P$  is the proportion of case vs. control and  $z = e^{-T^2/2}/\sqrt{2\pi}$  is the height of the standard normal probability density function at the truncation threshold  $T = \Phi^{-1}(1-K)$ .

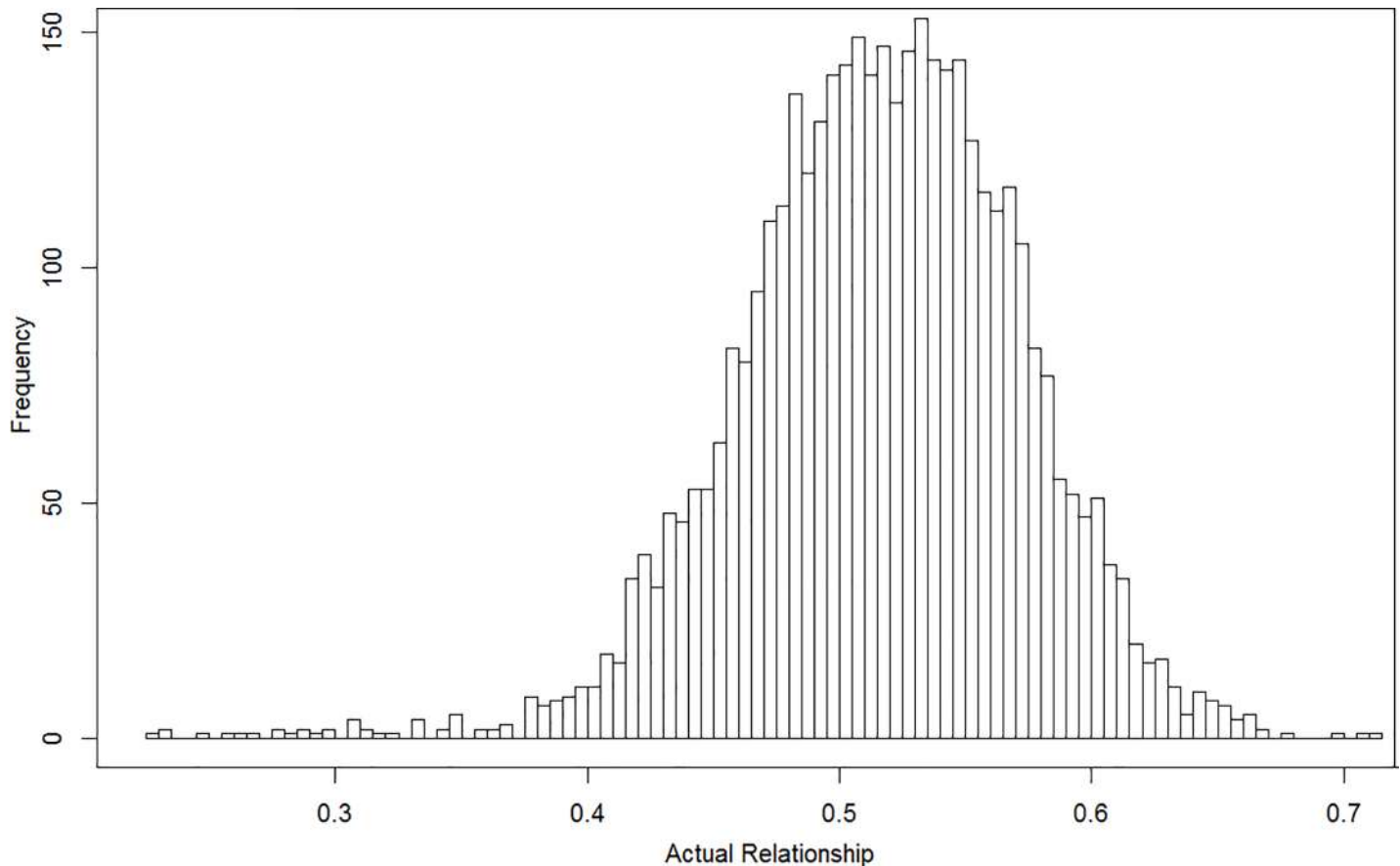
## Results

### IBD estimate

IBD probabilities were calculated using Plink [21]. Fig 1 shows the distribution of the genome-wide additive coefficients. The average proportion of the genome-shared IBD between the sib pairs (the coefficient of additive genetic variance) was 0.516 (standard deviation 0.056), with a range of 0.226–0.715.

### Heritability

We used a non-parametric bootstrap method [22] to calculate the standard error of the heritability. We divided the sib-pair dataset into three subsets, namely affected-affected sib-pairs,



**Fig 1. Histogram of the genome-wide additive genetic relationships of full-sib pairs estimated from genetic markers using Plink.**

<https://doi.org/10.1371/journal.pone.0176341.g001>

affected-unaffected sib-pairs and unaffected-unaffected sib-pairs. From each subset, we randomly selected with replacement 300 sib-pairs to reconstruct a new dataset of 900 samples where the proportion of case vs. control was always fixed at 0.5. We repeated this bootstrap routine 10,000 times to generate 10,000 new different datasets. Table 1 shows the overall  $h^2_L$  of T1D ranging from 0.53 to 0.92 depends on different settings of the T1D prevalence  $K$ . Our heritability estimates using sib-pair methods are closed to the results obtained from twin studies, ranging from 0.4 to 0.88 [5–8]. The R program and IBD data are available in Supporting Information files.

### AUC

The AUC and the corresponding 95% CIs for the sib-pair logistic regression model obtained in the discovery and validation sets are shown in Table 2. The AUC for the model was 0.75

**Table 1. The heritability on liability scale ( $h^2_L$ ) of T1D estimates using the well-known Haseman-Elston regression analysis.**

T1D prevalence $K$	$h^2_L$ (s.e.)
0.003 (general population)	0.53 (0.0017)
0.005 (general population)	0.59 (0.0019)
0.01 (siblings of affected probands)	0.69 (0.0022)
0.03 (siblings of affected probands)	0.92 (0.0029)

<https://doi.org/10.1371/journal.pone.0176341.t001>

**Table 2. AUC on T1D discovery and validation datasets generated from logistic regression models.**

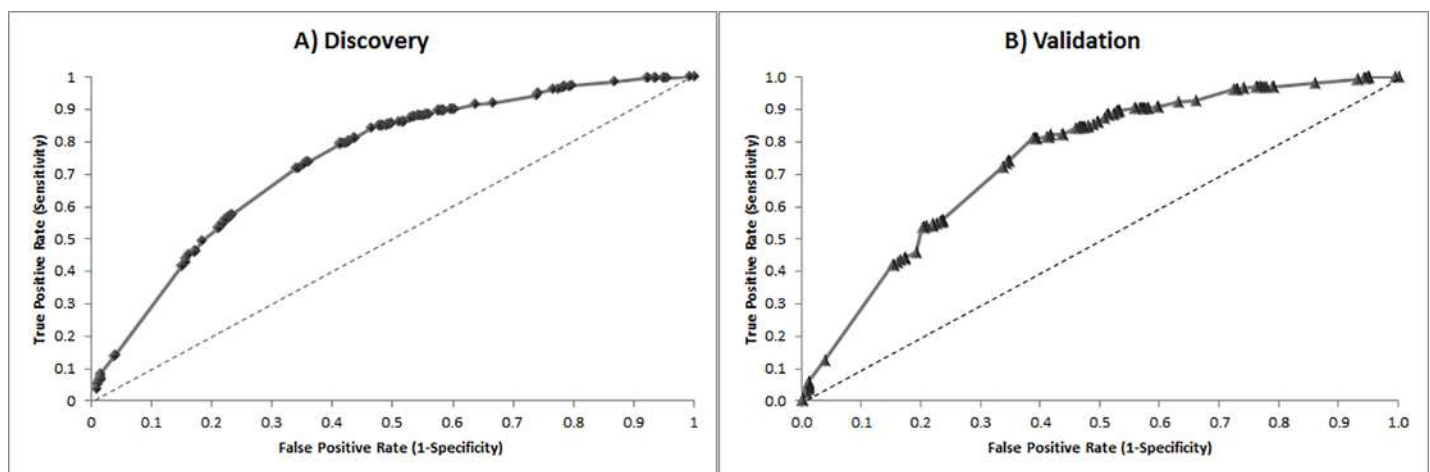
Model	Discovery set n = 750 pairs (1,500 samples)	Validation set N = 503 pairs (1,006 samples)
Presence / Absence DR3/4	0.61 (s.e. 0.014) 95% CI (0.59–0.64)	0.61 (s.e. 0.018) 95% CI (0.57–0.64)
3 SNPs <i>rs2854275</i> , <i>rs3104413</i> , <i>rs9273363</i>	0.75 (s.e. 0.013) 95% C.I. (0.72–0.77)	0.75 (s.e. 0.015) 95% C.I. (0.72–0.78)

<https://doi.org/10.1371/journal.pone.0176341.t002>

(95% CI, 0.72–0.77) in the discovery set, and when applied to the validation set, the AUC was also 0.75 (95% CI, 0.72–0.78). Thus, the model revealed consistency between the discovery and the validation sets. The overall AUCs in both datasets are far better than those of the model using only the presence / absence of the high risk HLA-DR3/4, namely AUC of 0.61 (s.e. 0.014–0.018). Fig 2 shows the ROC analyses from the sib-pair logistic regression model applied on the two datasets.

### Discussion

In the past 15 years, the genetics of common human diseases has been transformed by GWAS. These studies have been a powerful approach to the identification of genes involved in these complex diseases and led to developing predictive genetic tests. The tests using SNPs to predict an individual’s future risk of disease are one of the most appealing early disease prediction methods. Such tests can be conducted at birth and, by use of appropriate prevention strategies, prevent individuals from contracting a disease. These tests have the potential to be the cornerstone of epidemiology and are anticipated to have a large impact on health care (see further reviews in refs [23–25]). It is important to note that ROC curve is a simple and convenient overall measure of diagnostic test accuracy and does not depend on the prevalence of disease in the actual population. However, to measure the performance of a prediction model in clinical settings, the positive predictive value (PPV) and negative predictive value (NPV), which incorporate the disease prevalence in the testing population, are necessary. PPV is the proportion of patients who test positive for the disease who actually have the disease, and the NPV is the proportion of subjects who test negative who are actually free of the disease. Note that, like sensitivity and specificity, the positive and negative predictive values are dependent on



**Fig 2. ROC analyses from logistic regression models on A) T1D’s discovery set n = 750 pairs (1,500 samples) and B) T1D’s validation dataset n = 503 pairs (1,006 samples). Each point represents a test defined by a different logit score.**

<https://doi.org/10.1371/journal.pone.0176341.g002>

the risk score threshold  $T$  (the logit score in this study). When disease is rare like T1D, the threshold should be selected toward lower left portion of the ROC curve where the sensitivity is small but the specificity is high [13]. For example, when siblings of affected probands are screened, the proposed 3-SNP model achieves the PPVs of 2.7% and 7.8%, and the NPVs of 99.3% and 97.9% for the prevalence of 1% and 3%, respectively. In this case, screening for low prevalence T1D disease is cost effective because the cost of screening is less than the cost of care if the disease is not detected before disease onset. T1D, an autoimmune disease resulting from immune-mediated destruction of the insulin-producing  $\beta$ -islet cells of the pancreas, causes substantial morbidity and mortality and requires life-long insulin treatment. By the time that the disease is detected clinically, the  $\beta$ -cells are almost completely destroyed, and no known treatment can restore them.

Even though preventing or curing type 1 diabetes at risk subjects remain elusive despite effortless and substantial investments in industrialized countries [26,27], diabetes prevention research has been developed rapidly in recent years. In addition to current impressive methods to prevent type 1 diabetes such as metabolic modifications, antigen-specific vaccination, pancreatic transplantation, stimulation of  $\beta$ -cell regeneration, or avoidance of environmental triggers of islet autoimmunity [28], advances in stem cell biology, cell encapsulation methodologies, and immunotherapy will benefit the lives of patients in the end [27,29,30]. Importantly, if early onset diabetes of young children could be identified, screening high risk individuals can stave off or even avoid the short term as well as long term complications of type 1 diabetes. For short term complications, monitoring sugar can prevent new-onset diabetic ketoacidosis, the most severe acute diabetes-related central nervous system complication in young patients [31]. Early monitoring and modification of insulin sensitivity can also hamper diabetic nephropathy, one of the major causes of morbidity and mortality in type 1 diabetes [32,33]. The mortality and morbidity of heart disease are significantly escalated in type 1 diabetes patients compared to the nondiabetic population. An intervention at early stage to achieve glycaemia as close to normal as possible could alleviate and/or delay all of the cardiovascular complications of diabetes in high risk patients [34,35]. Several international projects for diabetes prevention such as DIPP [36], TEDDY [37], TRIGR [38], TrialNet [39] have screened and monitored thousands of newborn infants for HLA-DQB1 allele association with susceptibility to type 1 diabetes. As shown in the results, our proposed method is more accurate and much cheaper than the typical HLA typing.

Genetic factors play a significant role in T1D disease, as indicated by the proportion of explained variance ( $h^2$ ). As the heritability estimates for T1D explain ~90% of the phenotypic variance the GWAS-based predictions can be significantly improved by incorporating many other factors. These include invoking rare variants, structural variants, interaction between genes and environment factors, non-linear interaction between genes and genes, family history conditional on genotype at known loci and signals in non-coding regions [20,23–25].

## Supporting information

**S1 File. Histogram R.** R program for producing histogram of IBD data.

(R)

**S2 File. Heritability R.** R program for calculating t1d h2 using Haseman-Elston regression analysis.

(R)

**S3 File. Sibling IBD data.** Plink IBD probabilities for all sib-pairs studied in this project.

(TXT)

**S4 File. Sibling IBD data.** Plink IBD probabilities for non-affected vs non-affected sib pairs. (TXT)

**S5 File. Sibling IBD data.** Plink IBD probabilities for non-affected vs affected sib pairs. (TXT)

**S6 File. Sibling IBD data.** Plink IBD probabilities for affected vs affected sib pairs. (TXT)

## Acknowledgments

We are much grateful to Dr. Grant Morahan, Centre for Diabetes Research, University of Western Australia, for his invaluable comments and insightful input and we sincerely thank the anonymous reviewers for their time and effort into improving our manuscript.

## Author Contributions

**Conceptualization:** CN HL.

**Data curation:** CN.

**Formal analysis:** HL DN CN.

**Investigation:** HL DN CN.

**Methodology:** HL DN CN.

**Project administration:** HL DN.

**Resources:** DN CN.

**Software:** DN CN.

**Supervision:** HL CN.

**Validation:** DN.

**Visualization:** DN.

**Writing – original draft:** HL DN CN.

**Writing – review & editing:** HL DN CN.

## References

1. Jakobsdottir J, Gorin MB, Conley YP, Ferrell RE, Weeks DE. Interpretation of genetic association studies: markers with replicated highly significant odds ratios may be poor classifiers. *PLoS Genet.* 2009 Feb 6; 5(2):e1000337. <https://doi.org/10.1371/journal.pgen.1000337> PMID: 19197355
2. Morahan G, Varney M. The Genetics of Type 1 Diabetes. In: *The HLA Complex in Biology and Medicine A Resource Book*. New Delhi: JayPee Brothers Publishing; 2010:205–18.
3. Morahan G. Insights into type 1 diabetes provided by genetic analyses. *Current Opinion in Endocrinology, Diabetes and Obesity.* 2012 Aug 1; 19(4):263–70.
4. Wei Z, Wang K, Qu HQ, Zhang H, Bradfield J, Kim C, et al. From disease association to risk assessment: an optimistic view from genome-wide association studies on type 1 diabetes. *PLoS Genet.* 2009 Oct 9; 5(10):e1000678. Reich T, James JW, Morris CA. The use of multiple thresholds in determining the mode of transmission of semi-continuous traits. *Annals of human genetics.* 1972 Nov 1; 36(2):163–84. <https://doi.org/10.1371/journal.pgen.1000678> PMID: 19816555
5. Reich T, James JW, Morris CA. The use of multiple thresholds in determining the mode of transmission of semi-continuous traits. *Annals of human genetics.* 1972 Nov 1; 36(2):163–84 PMID: 4676360



6. Kyvik KO, Green A, Beck-Nielsen H. Concordance rates of insulin dependent diabetes mellitus: a population based study of young Danish twins. *Bmj*. 1995 Oct 7; 311(7010):913–7. PMID: [7580548](#)
7. Kaprio J, Tuomilehto J, Koskenvuo M, Romanov K, Reunanen A, Eriksson J, et al. Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. *Diabetologia*. 1992 Nov 1; 35(11):1060–7. PMID: [1473616](#)
8. Hyttinen V, Kaprio J, Kinnunen L, Koskenvuo M, Tuomilehto J. Genetic liability of type 1 diabetes and the onset age among 22,650 young Finnish twin pairs. *Diabetes*. 2003 Apr 1; 52(4):1052–5. PMID: [12663480](#)
9. Atkinson MA, Eisenbarth GS. Type 1 diabetes: new perspectives on disease pathogenesis and treatment. *The Lancet*. 2001 Jul 21; 358(9277):221–9.
10. Van Belle TL, Coppieters KT, Von Herrath MG. Type 1 diabetes: etiology, immunology, and therapeutic strategies. *Physiological reviews*. 2011 Jan 1; 91(1):79–118. <https://doi.org/10.1152/physrev.00003.2010> PMID: [21248163](#)
11. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nature genetics*. 2009 Jun 1; 41(6):703–7. <https://doi.org/10.1038/ng.381> PMID: [19430480](#)
12. Nguyen C, Varney MD, Harrison LC, Morahan G. Definition of high-risk type 1 diabetes HLA-DR and HLA-DQ types using only three single nucleotide polymorphisms. *Diabetes*. 2013 Jun 1; 62(6):2135–40. <https://doi.org/10.2337/db12-1398> PMID: [23378606](#)
13. Hosmer DW Jr, Lemeshow S, Sturdivant RX. *Applied logistic regression*. John Wiley & Sons; 2013 Apr 1.
14. Metz CE. Basic principles of ROC analysis. In *Seminars in nuclear medicine* 1978 Oct 1 (Vol. 8, No. 4, pp. 283–298). WB Saunders.
15. Fawcett T. An introduction to ROC analysis. *Pattern recognition letters*. 2006 Jun 30; 27(8):861–74.
16. Lu Q, Obuchowski N, Won S, Zhu X, Elston RC. Using the optimal robust receiver operating characteristic (ROC) curve for predictive genetic tests. *Biometrics*. 2010 Jun 1; 66(2):586–93. <https://doi.org/10.1111/j.1541-0420.2009.01278.x> PMID: [19508241](#)
17. Falconer, DS, Mackay TFC. *Introduction to Quantitative Genetics*. Addison 123 (Wesley Longman Ltd, 1996).
18. Haseman JK, Elston RC. The investigation of linkage between a quantitative trait and a marker locus. *Behavior genetics*. 1972 Mar 1; 2(1):3–19. PMID: [4157472](#)
19. Bates D, Maechler M, Bolker B, Walker S. *lme4: Linear mixed-effects models using Eigen and S4*. R package version. 2014 Jan;1(7).
20. Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. *The American Journal of Human Genetics*. 2011 Mar 11; 88(3):294–305. <https://doi.org/10.1016/j.ajhg.2011.02.002> PMID: [21376301](#)
21. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*. 2007 Sep 30; 81(3):559–75. <https://doi.org/10.1086/519795> PMID: [17701901](#)
22. Efron B. Bootstrap methods: another look at the jackknife. In *Breakthroughs in Statistics* 1992 (pp. 569–593). Springer New York.
23. Epstein CJ. Medical genetics in the genomic medicine of the 21st century. *The American Journal of Human Genetics*. 2006 Sep 30; 79(3):434–8. <https://doi.org/10.1086/507610> PMID: [16909381](#)
24. Manolio TA. Bringing genome-wide association findings into clinical use. *Nature Reviews Genetics*. 2013 Aug 1; 14(8):549–58. <https://doi.org/10.1038/nrg3523> PMID: [23835440](#)
25. Jostins L, Barrett JC. Genetic risk prediction in complex disease. *Human molecular genetics*. 2011 Oct 15; 20(R2):R182–8. <https://doi.org/10.1093/hmg/ddr378> PMID: [21873261](#)
26. Atkinson MA, Eisenbarth GS, Michels AW. Type 1 diabetes. *The Lancet*. 2014 Jan 10; 383(9911):69–82.
27. Ziegler AG, Bonifacio E, Powers AC, Todd JA, Harrison LC, Atkinson MA. Type 1 diabetes prevention: a goal dependent on accepting a diagnosis of an asymptomatic disease. *Diabetes*. 2016 Nov 1; 65(11):3233–9. <https://doi.org/10.2337/db16-0687> PMID: [27959859](#)
28. Gillespie KM. Type 1 diabetes: pathogenesis and prevention. *Canadian Medical Association Journal*. 2006 Jul 18; 175(2):165–70 <https://doi.org/10.1503/cmaj.060244> PMID: [16847277](#)
29. Bresson D, von Herrath M. Immunotherapy for the prevention and treatment of type 1 diabetes. *Diabetes care*. 2009 Oct 1; 32(10):1753–68 <https://doi.org/10.2337/dc09-0373> PMID: [19794001](#)
30. Vetere A, Choudhary A, Burns SM, Wagner BK. Targeting the pancreatic  $\beta$ -cell to treat diabetes. *Nature reviews Drug discovery*. 2014 Apr 1; 13(4):278–89. <https://doi.org/10.1038/nrd4231> PMID: [24525781](#)



31. Cameron FJ, Scratch SE, Nadebaum C, Northam EA, Koves I, Jennings J, et al. Neurological consequences of diabetic ketoacidosis at initial presentation of type 1 diabetes in a prospective cohort study of children. *Diabetes care*. 2014 Jun 1; 37(6):1554–62. <https://doi.org/10.2337/dc13-1904> PMID: 24855156
32. Bjornstad P, Cherney D, Maahs DM. Early diabetic nephropathy in type 1 diabetes—new insights. *Current opinion in endocrinology, diabetes, and obesity*. 2014 Aug; 21(4):279. <https://doi.org/10.1097/MED.000000000000074> PMID: 24983394
33. Bjornstad P, Snell-Bergeon JK, Rewers M, Jalal D, Chonchol MB, Johnson RJ, et al. Early Diabetic Nephropathy. *Diabetes Care*. 2013 Nov 1; 36(11):3678–83. <https://doi.org/10.2337/dc13-0631> PMID: 24026551
34. Nathan DM, DCCT/Edic Research Group. The diabetes control and complications trial/epidemiology of diabetes interventions and complications study at 30 years: overview. *Diabetes care*. 2014 Jan 1; 37(1):9–16 <https://doi.org/10.2337/dc13-2112> PMID: 24356592
35. Parikka V, Näntö-Salonen K, Saarinen M, Simell T, Ilonen J, Hyöty H, et al. Early seroconversion and rapidly increasing autoantibody concentrations predict prepubertal manifestation of type 1 diabetes in children at genetic risk. *Diabetologia*. 2012 Jul 1; 55(7):1926–36. <https://doi.org/10.1007/s00125-012-2523-3> PMID: 22441569
36. Virtanen SM, Kenward MG, Erkkola M, Kautiainen S, Kronberg-Kippilä C, Hakulinen T, et al. Age at introduction of new foods and advanced beta cell autoimmunity in young children with HLA-conferred susceptibility to type 1 diabetes. *Diabetologia*. 2006 Jul 1; 49(7):1512–21. <https://doi.org/10.1007/s00125-006-0236-1> PMID: 16596359
37. Hagopian WA, Lernmark Å, Rewers MJ, Simell OG, SheE JX, Ziegler AG, et al. TEDDY—the environmental determinants of diabetes in the young. *Annals of the New York Academy of Sciences*. 2006 Oct 1; 1079(1):320–6.
38. Krischer J, De Beaufort C. Study design of the Trial to Reduce IDDM in the Genetically Risk (TRIGR). *Pediatric diabetes*. 2007; 8:117–37 <https://doi.org/10.1111/j.1399-5448.2007.00239.x> PMID: 17550422
39. Skyler JS, Greenbaum CJ, Lachin JM, Leschek E, Rafkin-Mervis L, Savage P, et al. Type 1 Diabetes TrialNet—an international collaborative clinical trials network. *Annals of the New York Academy of Sciences*. 2008 Dec 1; 1150(1):14–24.