

Clinical Research Article

Prediction of Type 1 Diabetes at Birth: Cord Blood Metabolites vs Genetic Risk Score in the Norwegian Mother, Father, and Child Cohort

German Tapia,¹ Tommi Suvitaival,² Linda Ahonen,^{2,3} Nicolai A. Lund-Blix,^{1,4} Pål R. Njølstad,^{5,6} Geir Joner,^{4,7} Torild Skriverhaug,^{4,7} Cristina Legido-Quigley,² Ketil Størdal,^{1,4,8} and Lars C. Stene¹

¹Department of Chronic Diseases and Ageing, Norwegian Institute of Public Health, 0213 Oslo, Norway; ²Steno Diabetes Center Copenhagen, 2820 Gentofte, Denmark; ³Biosyntia ApS, 2100 Copenhagen, Denmark; ⁴Division of Pediatric and Adolescent Medicine, Oslo University Hospital, 0450 Oslo, Norway; ⁵Department of Pediatric and Adolescent Medicine, Haukeland University Hospital, 5021 Bergen, Norway; ⁶Center for Diabetes Research, Department of Clinical Science, University of Bergen, 5021 Bergen, Norway; ⁷Institute of Clinical Medicine, University of Oslo, 0318 Oslo, Norway; and ⁸Pediatric Research Institute, Institute of Clinical Medicine University of Oslo, 0372 Oslo, Norway

ORCID numbers: 0000-0002-1770-1183 (G. Tapia); 0000-0002-2583-4912 (T. Suvitaival); 0000-0002-9092-7867 (N. A. Lund-Blix); 0000-0003-0304-6728 (P. R. Njølstad); 0000-0001-7556-8141 (G. Joner); 0000-0002-4352-5929 (T. Skriverhaug); 0000-0002-4018-214X (C. Legido-Quigley); 0000-0002-8434-119X (L. C. Stene).

Abbreviations: aOR, adjusted odds ratio; AUC, area under the receiver operating characteristic curve; BMI, body mass index; DAISY, Diabetes Autoimmunity Study in the Young; DIPP, Type 1 Diabetes Prediction and Prevention; GRS, genetic risk score; MoBa, Mother, Father, and Child Cohort Study; PCA, principal component analysis; SNP, single-nucleotide polymorphism; TEDDY, The Environmental Determinants of Diabetes in the Young

Received: 9 March 2021; Editorial Decision: 31 May 2021; First Published Online: 4 June 2021; Corrected and Typeset: 5 July 2021.

Abstract

Background and aim: Genetic markers are established as predictive of type 1 diabetes, but unknown early life environment is believed to be involved. Umbilical cord blood may reflect perinatal metabolism and exposures. We studied whether selected polar metabolites in cord blood contribute to prediction of type 1 diabetes.

Methods: Using a targeted UHPLC-QQQ-MS platform, we quantified 27 low-molecular-weight metabolites (including amino acids, small organic acids, and bile acids) in 166 children, who later developed type 1 diabetes, and 177 random control children in the Norwegian Mother, Father, and Child cohort. We analyzed the data using logistic regression (estimating odds ratios per SD [adjusted odds ratio (aOR)]), area under the receiver operating characteristic curve (AUC), and k-means clustering. Metabolites were

compared to a genetic risk score based on 51 established non-HLA single-nucleotide polymorphisms, and a 4-category HLA risk group.

Results: The strongest associations for metabolites were amino adipic acid (aOR = 1.15; 95% CI, 0.97-1.55), indoxyl sulfate (aOR = 1.15; 95% CI, 0.87-1.51), and tryptophan (aOR = 0.84; 95% CI, 0.65-1.10), with other aORs close to 1.0, and none significantly associated with type 1 diabetes. K-means clustering identified 6 clusters, none of which were associated with type 1 diabetes. Cross-validated AUC showed no predictive value of metabolites (AUC 0.49), whereas the non-HLA genetic risk score AUC was 0.56 and the HLA risk group AUC was 0.78.

Conclusions: In this large study, we found no support of a predictive role of cord blood concentrations of selected bile acids and other small polar metabolites in the development of type 1 diabetes.

Key Words: metabolomics, cord blood, type 1 diabetes, LC-MS, Norwegian mother, father and child cohort

Type 1 diabetes is usually preceded by a prodromal phase characterized by islet autoantibodies, often appearing in early childhood years before diagnosis (1). HLA and other genetic factors clearly contribute to type 1 diabetes susceptibility (1), but the increasing incidence implicate nongenetic factors (2). The typically early seroconversion to islet autoantibodies suggests that early life is important (3-5).

Maternal age, obesity, and birth weight are early life nongenetic risk factors relatively consistently associated with childhood-onset type 1 diabetes (6-11). Obesity, dysglycemia, kidney function, and related traits, both in nonpregnant and in pregnant women, are associated with perturbations in small metabolites such as amino acids, creatinine, and bile acids (12-15), many of which have also been associated with birth weight (12). Metabolites such as glucose, lipids, amino acids, and bile acids can cross the placenta, often bidirectionally, via free diffusion and placentally expressed transmembrane transporters (16-18). Many small metabolites in cord blood are thus correlated with maternal levels during the third trimester (12). For example, plasma creatinine, a marker of kidney function, largely reflects maternal levels when measured in cord blood (19), but has been linked to birth weight (20, 21). Maternal circulating bile acids, which may be influenced by maternal gut microbiota, may program offspring metabolism, or influence their microbiome, and have been linked to insulin resistance (14, 22, 23). Yet, there is only 1 small study (15 cases and 24 controls) in cord blood (24) and 1 study using dried blood spots (25) to date on nonlipid metabolites and later type 1 diabetes. The authors are not aware of any previous study investigating maternal or newborn plasma bile acids and subsequent type 1 diabetes risk.

The aim of the study was to test if selected small metabolites in cord blood, or combinations of these, could predict future risk of offspring type 1 diabetes in the Norwegian Mother, Father, and Child Cohort Study (MoBa), 1 of the largest pregnancy cohorts in the world. To ensure robust quantification, and to minimize multiple testing problems, we chose a targeted metabolomics approach focusing on small metabolites (molar mass ranging from 75 to ~500 g/mol) with previous evidence for association with metabolic traits (13). In addition, we investigated established genetic susceptibility markers for comparison of predictive values among biomarkers present at birth.

Materials and Methods

Participants and study design

We designed a nested case-control study in the MoBa cohort (26), which recruited ~114 000 pregnant mothers (41% eligible participated) nationwide from 1999 through 2008 (last birth in 2009). The current study uses data from cord blood samples and repeated questionnaires collected during pregnancy and up to child age 6 months (27). All participating mothers gave written informed consent. The establishment of MoBa and initial data collection was based on a license from the Norwegian Data Protection Agency and approval from The Regional Committees for Medical and Health Research Ethics. The MoBa cohort is now based on regulations related to the Norwegian Health Registry Act. The current study was approved by The Regional Committees for Medical and Health Research Ethics. Children who developed type 1 diabetes by February 5, 2014, were identified by register linkage to the Norwegian Childhood Diabetes Registry (28) and selected as cases. A random sample of the cohort was included as controls. Case and control samples were retrieved simultaneously

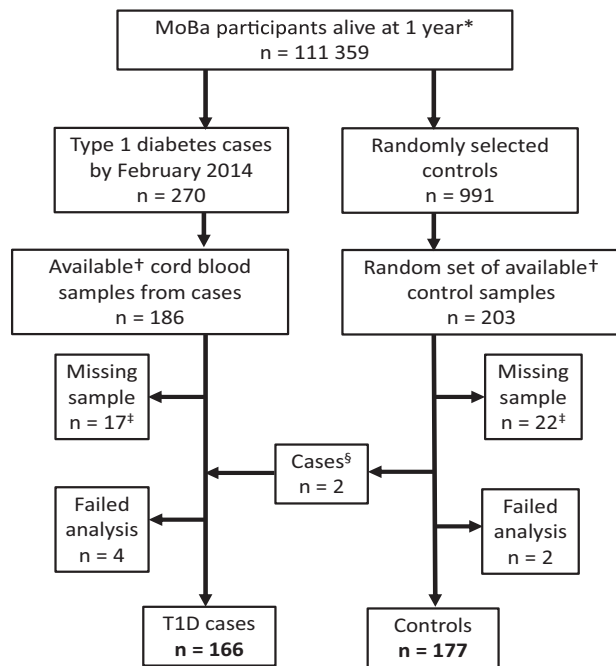


Figure 1. Flowchart of participants. *Live-born children who survived their first year of life. †Cord blood samples registered, processed, and not reserved for other projects. ‡Includes cord blood samples with too low volume and unseparated blood. §We ascertained if any random control child developed type 1 diabetes by June 1, 2018; these (n = 2) were reclassified as cases.

and treated equally. In total, 166 children were type 1 diabetes cases and 177 children were controls (Fig. 1). Baseline characteristics for those with available blood samples were largely similar to the whole MoBa cohort, except a lower proportion of cesarean section and preterm birth (29).

Cord blood metabolomics profiling

Blood was sampled from the umbilical cord vein using a syringe immediately after birth, collected in a EDTA container and shipped to the biobank for plasma separation and storage at -80°C (30). We used a UHPLC-QQQ-MS platform for targeted metabolomics with focus on robust measurement as described by Ahonen et al (13). Briefly, measurements were normalized, outliers were removed, and then metabolites were excluded on the basis of high missingness (n = 1) or low coefficient of variation (n = 3) before analysis. Details are given in the Supplemental Appendix (31). In total, 27 metabolites were included for analysis, representing various metabolic pathways. A brief description of the selected metabolites is provided in Supplemental Table 1 (31). Mean values and number of nonmissing measurements of the metabolites, in cases and controls, are shown in Supplemental Table 2 (31). The distributions of the

plasma metabolite concentrations were approximately symmetric after log-transformation (Supplemental Figure 1) (31). Z scores were calculated to compare metabolites across different scales.

Genotyping assays and genetic risk scores

Participants were genotyped for established type 1 diabetes susceptibility single-nucleotide polymorphisms (SNPs) including HLA tagSNPs using a custom Illumina Golden Gate assay (Illumina, San Diego, CA), as described earlier (32). We calculated a 51-SNP non-HLA type 1 diabetes genetic risk score (GRS), weighted by the natural log-odds ratio of type 1 diabetes per risk allele reported from large genome-wide association studies (Supplemental Table 3 (31), with further details given in the online supplement to (33)). HLA class II alleles were imputed using the HLA*IMP:02 web service (34) (details given previously (32)) and all HLA genotypes were subsequently confirmed by allele specific PCR (35).

Other covariates

Information on birth weight, maternal age at delivery, and delivery mode was obtained from the nationwide Medical Birth Registry of Norway (36). Information regarding maternal prepregnancy body mass index (BMI) and smoking during pregnancy was obtained from midpregnancy and child's 6-months-of-age questionnaires. Questionnaires can be accessed at www.fhi.no/moba. Maternal type 1 diabetes data were obtained from questionnaires and the Norwegian Patient Registry. Variables were categorized as shown in Table 1.

Statistical methods

We used logistic regression with child type 1 diabetes as outcome, adjusted for selected variables, unless noted otherwise. The primary and secondary analyses were decided a priori. Our primary analysis was to estimate the logit-linear association between each single cord blood metabolite and child type 1 diabetes risk. Tertiles of each metabolite were used as categorical exposures to test our linearity assumption. We generated 6 clusters (groups of individuals with similar patterns across all metabolites passing quality control) using k-means clustering, assessing whether the 6 clusters associated with type 1 diabetes risk, to investigate if any pattern of metabolites associated with type 1 diabetes.

Secondary analyses included restricting the primary analysis to children carrying type 1 diabetes risk HLA genotypes. We also evaluated the predictive performance of

all the metabolites simultaneously, by calculating the area under the receiver operating characteristic curve (AUC) and compared with the AUC of well-established genetic type 1 diabetes susceptibility markers. A model including z scores of all analyzed metabolites, a model using HLA and a model using a weighted GRS were used to generate predictions for later type 1 diabetes based on metabolites, HLA, and non-HLA GRS, respectively. Five-fold cross validation was used in the AUC analysis to adjust for potential overfitting. We also investigated groups of metabolites by using the sum of amino acid levels, sum of bile acid levels, and the ratio of taurine/glycine-conjugated bile acids as exposures.

Exploratory analyses included principal component analysis (PCA) using principal components with an eigenvalue above 1 as predictors, and the MetaboAnalyst web service

for metabolite set enrichment analysis (37). Details of the statistical analysis are given in Supplemental Methods (31).

Missing values were excluded in the analyses, with the exception of the cluster, AUC, PCA, enrichment, and groups (amino and bile acids) and ratios of metabolites analyses, where we imputed the sample mean for missing values to allow inclusion of all metabolites. We used clustered sandwich estimator for standard errors to account for correlation between siblings. We present nominal P values (not corrected for multiple testing), and 95% CIs for the odds ratio excluding 1.00 were considered equivalent to a nominal significance at the 2-sided 5% level. We planned a priori to use false discovery rate and calculate q -values to account for multiple testing with 27 metabolites.

Table 1. Characteristics of participants

	Controls (n = 177)	Cases (n = 166)	<i>P</i>
Age, median (range) at end of follow-up, ^a y	10.8 (7.1-15.7)	12.0 (7.3-15.6) ^b	<0.01
Female sex	80 (45.2%)	81 (48.8%)	0.51
Preterm birth ^c	5 (2.8%)	10 (6.0%)	0.15
Cesarean section ^d	12 (6.8%)	28 (16.9%)	<0.01
Parity			0.42
No earlier births	77 (43.5%)	83 (50.0%)	
1	63 (35.6%)	49 (29.5%)	
2 or more	37 (20.9%)	34 (20.5%)	
Maternal age, y, median (range)	30 (20-42)	30 (19-42)	0.29
19-24	24 (13.6%)	14 (8.4%)	
25-34	128 (72.3%)	130 (78.3%)	
35-42	25 (14.1%)	22 (13.3%)	
Maternal smoking during pregnancy ^e			0.73
Nonsmoker at end of pregnancy	153 (86.4%)	143 (86.1%)	
Smoked at end of pregnancy	17 (9.6%)	18 (10.8%)	
Missing data	7 (4.0%)	5 (3.0%)	
Prepregnancy body mass index, kg/m ² , median (range)	22.8 (17.2-40)	23.9 (15.8-40.5)	<0.01
<25	123 (69.5%)	88 (53.0%)	
25-30	27 (15.3%)	42 (25.3%)	
>30	15 (8.5%)	24 (14.5%)	
Missing data	12 (6.8%)	12 (7.2%)	
Child's HLA genotype ^f			<0.01
Protective (DQ6)	55 (31.1%)	4 (2.4%)	
Neutral (any other HLA not mentioned)	29 (16.4%)	5 (3.0%)	
Increased risk (≥ 1 copy of either DQ8 or DQ2.5)	70 (39.5%)	86 (51.8%)	
High risk (DQ8/DQ2.5 heterozygote)	11 (6.2%)	67 (40.4%)	
Missing data	12 (6.8%)	4 (2.4%)	

P values from Pearson's χ^2 test.

^aEnd of follow-up was set as the last diagnosis date of the cases on June 2, 2016.

^bMedian age of diagnosis was 5.8 (range, 0.7-12.8); only 3 cases were diagnosed after 12 years.

^cDefined as birth before pregnancy week 37.

^dIncludes elective (7 controls, 12 cases) and acute (5 controls, 16 cases) cesarean sections.

^eIncluding those that quit smoking shortly before or during pregnancy because the protective association with type 1 diabetes has been observed in those that smoked throughout pregnancy (49).

^fCoded as a binary variable in the analysis—protective or neutral (0) vs increased or high risk (1). The increased risk category (≥ 1 copy of either DQ8 [DQA1*03-DQB1*03:02-DRB1*04] or DQ2.5 [DQA1*05:01-DQB1*02:01] also required no HLA DQ6 [DQA1*01:02-DQB1*06:02-DRB1*15:01] protective allele).

The following covariates were included in our primary adjustment model: sample batch, date of run, child's sex, cesarean delivery, gestational age at birth (in weeks), maternal age, prepregnancy BMI, smoking in pregnancy, and parity (categorized as shown in Table 1). Analyses were run in R version 4.0.2 (38) and Stata release 16 (Stata Corp LLC, College Station, TX).

Results

Pairwise correlation coefficients among metabolites ranged from 0.77 for glycocholate-taurocholate and 0.62 for tyrosine-phenylalanine, to inverse correlations of -0.74 for glutamine-glutamic acid and -0.55 for taurine-glutamine, but most metabolite pairs showed little correlation (Supplemental Figure 2) (31).

Main analysis: Association of individual metabolites and clusters with type 1 diabetes

The strongest associations were observed for amino adipic acid (adjusted odds ratio [aOR] = 1.23; 95% CI, 0.97-1.55), indoxyl sulfate (aOR = 1.15; 95% CI, 0.87-1.51), taurochenodeoxycholic acid (aOR = 1.12; 95% CI, 0.87-1.44), and tryptophan (aOR = 0.84; 95% CI, 0.65-1.10), with other aORs close to 1.0. No metabolite was

nominally significantly associated with type 1 diabetes (Fig. 2). Categorizing metabolites into tertiles did not show deviation from linearity, except for phenylalanine (Supplemental Figure 3) (31). The second tertile of phenylalanine was associated with lower risk of type 1 diabetes, whereas the lowest and highest tertile had neutral risk (Supplemental Figure 3) (31), meaning the linear estimate reported must be cautiously interpreted. Adjusting for batch, date of run, and maternal BMI or birthweight only, variables that could be the most relevant for metabolite levels and offspring type 1 diabetes did not appreciably change our results (data not shown). When individuals were clustered using k-means clustering into 6 groups of cord blood metabolite profiles, no cluster was significantly associated with type 1 diabetes (Fig. 2).

Secondary analyses: Multivariable prediction of type 1 diabetes at birth with receiver operating characteristic curves

Using all 27 metabolites to predict later type 1 diabetes with multiple logistic regression gave an AUC of 0.66, the non-HLA genetic risk score alone an AUC of 0.57, and HLA risk group alone gave an AUC of 0.81, and all combined an AUC of 0.87 before cross validation (Fig. 3A). Because the non-HLA SNP was externally weighted, it was not expected to give a too optimistic prediction

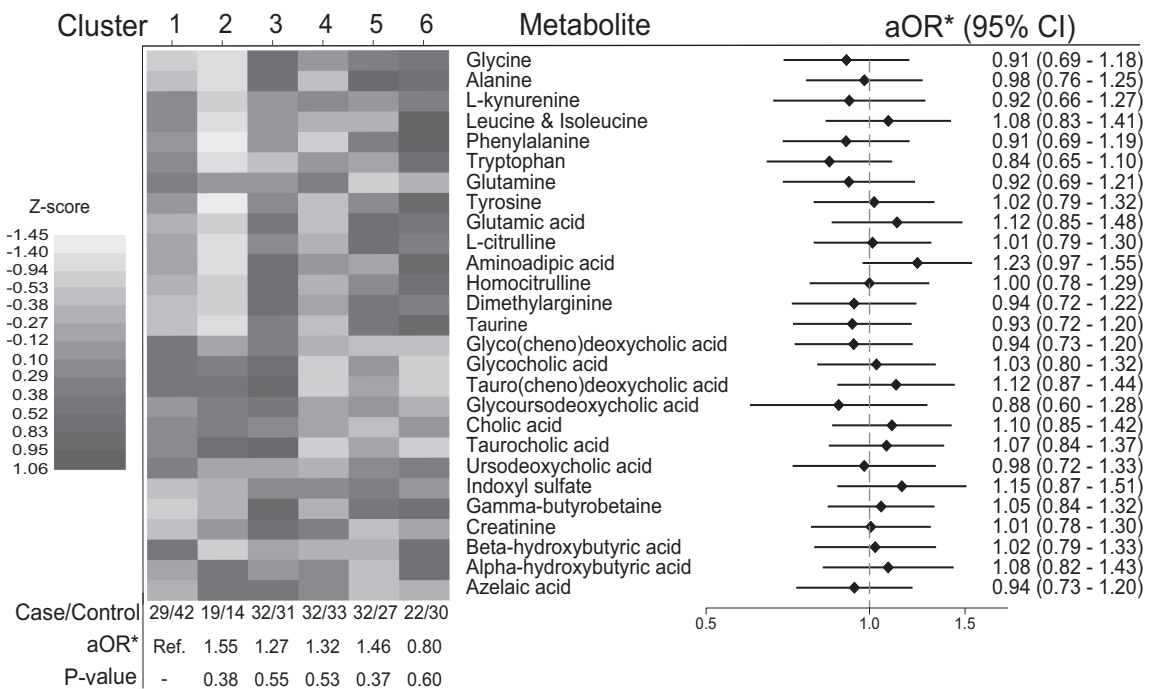


Figure 2. Shows the results from the cluster analysis (with a heatmap representing mean metabolite concentrations in each cluster) and the results from the main analysis (logistic regression, with each metabolite in a separate model) as a forest plot. *Adjusted odds ratio (aOR) for sample batch, date of run, child's sex, cesarean delivery, length of pregnancy (in weeks), maternal age, prepregnancy body mass index (BMI), smoking in pregnancy, and parity.

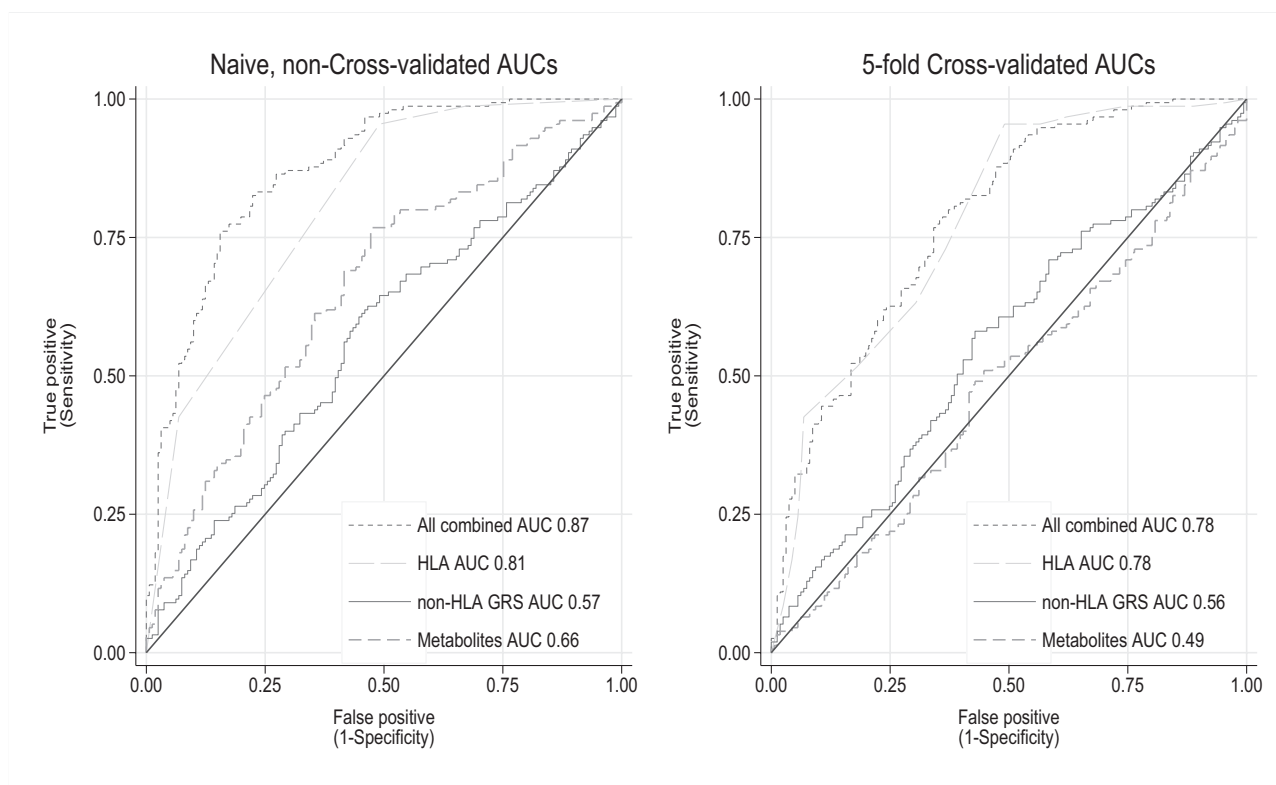


Figure 3. Area under the receiver operating characteristics curves (AUC) of predicted risk of type 1 diabetes, using predictions from logistic regression with either 27 cord blood metabolites, a 51 single-nucleotide polymorphism (SNP) non-HLA genetic risk score (GRS), 4-category HLA risk group, or all combined, both cross-validated (CV) and not.

(overfitting), whereas the HLA and metabolite predictions were not externally weighted. After accounting for potential overfitting using 5-fold cross-validation, the metabolites showed no predictive value, whereas the HLA group and non-HLA genetic risk score were only marginally attenuated after cross validation. HLA remained the clearly most important predictor (Fig. 3B). This illustrates that the initial suggestive predictive utility of metabolites was likely because of overfitting.

Other secondary analyses

Analyzing metabolite groups (by summing the z scores of amino acids, or bile acids, and using these variables as exposures) did not show any statistical association with later type 1 diabetes (amino acids aOR = 1.00; 95% CI, 0.96-1.03; bile acids aOR = 1.02; 95% CI, 0.95-1.09). The taurine/glycine-conjugated bile acid ratio was not associated with offspring type 1 diabetes (aOR = 1.07; 95% CI, 0.91-1.26). Restricting our analysis to only children carrying HLA genotypes conferring increased type 1 diabetes risk gave similar results as the main analysis (Supplemental Table 4) (31).

Exploratory analyses

Metabolite set enrichment analysis did not show any statistically significant results (Supplemental Figure 4) (31). PCA analysis of the metabolites detected 8 principal components with eigenvalues larger than 1. The cumulative proportion of variance of the components was 0.62, and none of these components was associated with later type 1 diabetes (Supplemental Figure 5) (31). Pairwise plotting of the principal components did not show any clear pattern (an example is shown in Supplemental Figure 6) (31).

Discussion

In this study, we investigated selected small polar metabolites measured in cord blood and their potential association with later type 1 diabetes. No metabolites, or cluster of individuals with different metabolite profile, were significantly associated with later type 1 diabetes.

Comparison to other studies

To our knowledge, there are no previous studies on maternal or infant bile acids and their potential association with type

1 diabetes. There are only 2 studies on polar metabolites at birth and type 1 diabetes. Orešič et al. reported lower levels of several metabolites, including tryptophan, glycine, and creatinine, in cord blood associated with later type 1 diabetes in the Finnish Type 1 Diabetes Prediction and Prevention (DIPP) cohort (24). La Marca et al reported lower levels of alanine and carnitines in children who later developed type 1 diabetes compared with children who did not using dried blood spots taken shortly after birth (25). Our findings are not consistent with results from these studies, but we did not measure carnitines. Cord blood and dried blood spots might also not be directly comparable.

Metabolomics during infancy and early childhood and association with islet autoimmunity or type 1 diabetes have been investigated in a few previous studies. The DIPP cohort reported lower glutamic acid and tryptophan at 3 or 6 months of age (q -values < 0.1) (39), and lower concentrations of branched chained amino acids, phenylalanine and tyrosine in peripheral blood mononuclear cells at 12 months of age (not significant after multiple testing correction), in children who developed type 1 diabetes compared with controls (40). In the international The Environmental Determinants of Diabetes in the Young (TEDDY) cohort, Johnson et al measured metabolites at 9 months of age and reported dicarboxylic acids (adipic acid being most significant) to be associated with increased risk of subsequent islet autoimmunity (41). Li et al, also based on the TEDDY cohort, reported alanine and β -hydroxybutyric acid at 12 and 24 months of age, respectively, to be associated with lower risk of developing islet autoimmunity in those developing autoantibodies against insulin first (42). The Norwegian Environmental Triggers of Type 1 Diabetes study reported differences in amino acid levels, albeit not statistically significant after correction for multiple testing (43), and the German BABYDIAB study reported lower methionine levels (44), after development of islet autoantibodies. There are also 2 publications from TEDDY and 1 from the Diabetes Autoimmunity Study in the Young (DAISY) using polar metabolites to predict later islet autoimmunity. Stanfill et al used classification algorithms to determine the most predictive features of islet autoimmunity in the TEDDY study (504 samples) and reported adipic acid, creatinine, and leucine as influential metabolites (45). Webb-Robertson et al used a machine learning approach to predict islet autoimmunity using data from the TEDDY study (157 case-control pairs), reporting azelaic acid and adipic acid as important features (46). Frohnert et al predicted seroconversion to islet autoimmunity in the DAISY study (22 cases and 25 controls) and reported 3-methyl-oxobutyrate (a precursor to valine for leucine synthesis) and pyroglutamate acid (a derivative of glutamic acid) as features that were often selected by the algorithm used (47).

We did not find any strong associations between specific metabolites measured at birth and later type 1 diabetes. Likewise, we did not find any cluster or group of metabolites that were overrepresented in children who later developed type 1 diabetes, and the metabolites studied had low predictive value for later type 1 diabetes. It must be kept in mind that earlier studies are heterogenous in most aspects: sample handling, measurement method, exposures measured, results reported, statistical analysis, and endpoint (for an overview, see Supplemental Table 5) (31), which makes results difficult to compare. Few studies adjusted for covariates but adjusting for covariates or stratifying by type 1 diabetes HLA risk did not change our estimates to a large degree. Approaches to control multiplicity varies and it is not always clear how the problem has been handled in previous studies. It is not known if participants/samples included in subsequent publications from the same cohort (such as publications from the DIPP and TEDDY cohorts) overlap to a large degree or are separate. With the exception of Orešič et al (24) and la Marca et al (25), other studies have not measured polar metabolites at birth. There could also be differences between countries or the population studied that lead to different results.

The previously published prediction studies using postnatal circulating metabolites are likewise not directly comparable, and there are few consistent findings. The reported AUCs differ across studies, as does the data analysis and validation approach. Frohnert et al, starting with 1552 features (and 22 cases), reported an AUC of 0.92 for prediction of islet autoimmunity in the DAISY study, without a holdout validation set (47). Webb-Robertson et al, using the TEDDY cohort, reported that the top 42 features (of 221) gave an AUC of 0.65 for prediction of islet autoimmunity in the holdout validation set, appreciably lower than the AUC of 0.74 estimated with cross-validation of the training set (46). As seen in our results, cross-validation is essential to avoid overfitting, even when using few features and a relatively large sample size (166 cases, 177 controls). Several factors influence the potential for overfitting and bias, but in general, the larger the number of predictors in relation to the number of subjects with outcome, and the more flexible modelling approach (allowing nonlinearities and interactions), the larger the potential for overfitting. Using a nonvalidated AUC analysis could lead to the misinterpretation that these metabolites have predictive value for later type 1 diabetes in the same range as well-established genetic risk factors, and would add to the predictive value of genetic factors. Cross-validation shows that this overly optimistic results is due to overfitting even with a simple logistic regression model, whereas well-established genetic factors had an essentially unchanged

AUC. This is consistent with the fact that overfitting bias tend to be smaller the greater the a priori evidence for the predictors used.

Although our approach was to start with a limited number of predictors ($n = 27$) in the model without statistical selection, an approach using statistical selection, such as stepwise selection or least absolute shrinkage and selection operator, would not select any metabolites and thus not obtain an overfitted model. We nevertheless think our results provide an important message regarding potential bias in multivariable prediction of type 1 diabetes or other outcomes. Although overfitting and multiple testing are related concepts, reducing the influence of these requires different techniques. Overfitting is a problem of multivariable prediction and can be minimized by cross-validation and independent validation data. For correcting multiple statistical testing, we had planned to control the false discovery rate at 5%, and compute q-values. Because no single metabolite was significant at the nominal 5% level, we conclude that no metabolite was significant, and q-values do not make any sense in this case (48).

Interpretation and implication of results

Our goal was to investigate if metabolites at birth could be predictive of later type 1 diabetes by studying selected metabolites both singly and combined and investigating their predictive value. Our interpretation of the data is that these metabolites, measured in cord blood, are of limited importance in offspring type 1 diabetes. The evidence that intrauterine exposures are associated with type 1 diabetes is not extensive (3). This does not mean that no metabolic changes occur closer to development of disease, but these possible differences do not seem present at birth for the measured metabolites. Lack of predictive value of these metabolites also does not negate the existence of other perinatal exposures that may influence the risk of type 1 diabetes, if we assume such factors are not strongly correlated with the metabolites measured in our study. The relatively consistently replicated associations between maternal obesity, birth weight, and type 1 diabetes could exert their potential effects through other means.

Strengths and limitations

A strength of the study is the use of cord blood samples, which allows us to study exposures before onset of autoimmunity and dysglycemia, and characterize influences during intrauterine life, including potential maternal exposures in pregnancy. We include children from

the general population, not selected by family history or HLA risk genotypes, which makes our results more applicable to the general population. This study is the largest measuring polar metabolites in cord blood. The targeted approach allowed us to quantify a set of metabolites with a biological basis. This reduced problems of multiplicity (multiple testing and overfitting), with the potential cost of not including potentially relevant metabolites. Still, increasing the number of measured metabolites would come at a cost of increasing multiplicity problems. As with any observational study, we cannot rule out unmeasured confounders.

Conclusions

In this large study, the selected polar metabolites measured in cord blood were not associated with later type 1 diabetes in the offspring.

Acknowledgments

The Norwegian Mother, Father, and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. We are grateful to all the participating families in Norway who take part in this ongoing cohort study. The Norwegian Childhood Diabetes Registry is funded by The South-Eastern Norway Regional Health Authority. We are thankful to the Norwegian Childhood Diabetes Registry and Study Group. Data from the Norwegian Patient Register have been used in this publication. The interpretation and reporting of these data are the sole responsibility of the authors, and no endorsement by the Norwegian Patient Register is intended nor should be inferred. Genotyping in this study was performed at the Genomics Core Facility, Norwegian Radium Hospital, Oslo, Norway. We are thankful to K. Gillespie and G. Mortimer (Diabetes and Metabolism, School of Clinical Sciences, Southmead Hospital, University of Bristol, Bristol, UK) for the HLA typing. In addition, the authors thank laboratory technician Birgitte Nergaard Roberts (Steno Diabetes Center Copenhagen), who was responsible for the extensive sample randomization and sample preparation before the UHPLC-QQQ-MS analyses.

Financial Support: Costs of all data acquisition, including laboratory assays in MoBa (the substudy PAGE; Prediction of Autoimmune Diabetes and Celiac Disease in Childhood by Genes and Perinatal Environment), was supported by grant 2210909/F20 from the Norwegian Research Council (L.C.S.). Supported by grants (to P.R.N.) from the European Research Council (AdG #293574), the Bergen Research Foundation (“Utilizing the Mother and Child Cohort and the Medical Birth Registry for Better Health”), Stiftelsen Kristian Gerhard Jebsen (Translational Medical Center), the University of Bergen, the Research Council of Norway (FRIPRO grant #240413), the Western Norway Regional Health Authority (Strategic Fund “Personalized Medicine for Children and Adults”), the Novo Nordisk Foundation (grant #54741), and the Norwegian Diabetes Association. This research was funded, in whole or in part, by The Research Council of Norway. For the purpose of open access, the

author has applied a CC BY public copyright license to any Author Accepted Manuscript (AAM) version arising from this submission.

Author Contributions: Conception and design: G.T., L.C.S., K.S. Literature search: G.T., L.C.S. Acquisition of pregnancy cohort data: L.C.S., K.S. Acquisition of incident type 1 diabetes data: T. Skrivarhaug, G.J., P.R.N. Measurement of cord blood metabolites: T. Suvisaival, L.A., C.L.Q. Data cleaning and preparation: G.T., T. Suvisaival. Planning statistical analyses: L.C.S., G.T., K.S. Performing statistical analyses: G.T. Interpretation of data: all authors. Drafting the manuscript: G.T. Revising the manuscript critically for important intellectual content: all authors. Final approval of the version to be published: all authors. Taking responsibility for the integrity of the data and the accuracy of the data analysis: G.T., L.C.S. Obtaining funding: L.C.S., K.S.

Preprint version available at: <https://www.medrxiv.org/content/10.1101/2021.02.20.21252128v1.full-text>

Supplementary Appendix with 6 Figures, 5 Tables, available at: https://figshare.com/articles/journal_contribution/Supplementary_Appendix_Tapia_Cord_Blood_Metabolites_T1D_pdf/14566476

Correspondence: German Tapia, Department of Chronic Diseases and Ageing, Norwegian Institute of Public Health, PO Box 222 Skøyen, NO-0213 Oslo, Norway. Email: german.tapia@fhi.no.

Disclosures: The authors have nothing to disclose. No potential conflict of interest relevant to this article was reported. The authors alone are responsible for the content and writing of the paper.

Data Availability: Aggregated data are available from the authors upon reasonable request. The consent given by the participants does not open for storage of data on an individual level in repositories or journals. Access to individual level data sets requires an application, approval from The Regional Committee for Medical and Health Research Ethics in Norway, and an agreement with MoBa.

References

- Atkinson MA, Eisenbarth GS, Michels AW. Type 1 diabetes. *Lancet*. 2014;383(9911):69-82.
- Tuomilehto J, Ogle GD, Lund-Blix NA, Stene LC. Update on worldwide trends in occurrence of childhood type 1 diabetes in 2020. *Pediatr Endocrinol Rev*. 2020;17(Suppl 1):198-209.
- Stene LC, Gale EA. The prenatal environment and type 1 diabetes. *Diabetologia*. 2013;56(9):1888-1897.
- Norris JM, Johnson RK, Stene LC. Type 1 diabetes-early life origins and changing epidemiology. *Lancet Diabetes Endocrinol*. 2020;8(3):226-238.
- Rewers M, Hyöty H, Lernmark Å, et al.; TEDDY Study Group. The environmental determinants of diabetes in the young (TEDDY) study: 2018 update. *Curr Diab Rep*. 2018;18(12):136.
- Stene LC, Magnus P, Lie RT, Søvik O, Joner G. Maternal and paternal age at delivery, birth order, and risk of childhood onset type 1 diabetes: population based cohort study. *Bmj*. 2001;323(7309):369.
- Cardwell CR, Stene LC, Joner G, et al. Birthweight and the risk of childhood-onset type 1 diabetes: a meta-analysis of observational studies using individual patient data. *Diabetologia*. 2010;53(4):641-651.
- Cardwell CR, Stene LC, Joner G, et al. Maternal age at birth and childhood type 1 diabetes: a pooled analysis of 30 observational studies. *Diabetes*. 2010;59(2):486-494.
- Magnus MC, Olsen SF, Granstrom C, et al. Paternal and maternal obesity but not gestational weight gain is associated with type 1 diabetes. *Int J Epidemiol*. 2018;47(2):417-426.
- Hidayat K, Zou SY, Shi BM. The influence of maternal body mass index, maternal diabetes mellitus, and maternal smoking during pregnancy on the risk of childhood-onset type 1 diabetes mellitus in the offspring: systematic review and meta-analysis of observational studies. *Obes Rev*. 2019;20(8):1106-1120.
- Waernbaum I, Dahlquist G, Lind T. Perinatal risk factors for type 1 diabetes revisited: a population-based register study. *Diabetologia*. 2019;62(7):1173-1184.
- Lowe WL Jr, Bain JR, Nodzenski M, et al.; HAPO Study Cooperative Research Group. Maternal BMI and glycemia impact the fetal metabolome. *Diabetes Care*. 2017;40(7):902-910.
- Ahonen L, Jännti S, Suvisaival T, et al. Targeted clinical metabolite profiling platform for the stratification of diabetic patients. *Metabolites*. 2019;9(9):184.
- Haeusler RA, Astiarraga B, Camastra S, Accili D, Ferrannini E. Human insulin resistance is associated with increased plasma levels of 12 α -hydroxylated bile acids. *Diabetes*. 2013;62(12):4184-4191.
- Carter TC, Rein D, Padberg I, et al. Validation of a metabolite panel for early diagnosis of type 2 diabetes. *Metabolism*. 2016;65(9):1399-1408.
- Macias RI, Marin JJ, Serrano MA. Excretion of biliary compounds during intrauterine life. *World J Gastroenterol*. 2009;15(7):817-828.
- Carter AM. Evolution of placental function in mammals: the molecular basis of gas and nutrient transfer, hormone secretion, and immune responses. *Physiol Rev*. 2012;92(4):1543-1576.
- Michelsen TM, Holme AM, Henriksen T. Transplacental nutrient transfer in the human in vivo determined by 4 vessel sampling. *Placenta*. 2017;59 Suppl 1:S26-S31.
- Basalely A, Liu D, Kaskel FJ. Big equation for small kidneys: a newly proposed model to estimate neonatal GFR. *Pediatr Nephrol*. 2020;35(4):543-546.
- Lao TT, Loong EP, Chin RK, Lam YM. Renal function in the newborn. Newborn creatinine related to birth weight, maturity and maternal creatinine. *Gynecol Obstet Invest*. 1989;28(2):70-72.
- Morken NH, Travlos GS, Wilson RE, Eggesbø M, Longnecker MP. Maternal glomerular filtration rate in pregnancy and fetal size. *Plos One*. 2014;9(7):e101897.
- Papacleovoulou G, Abu-Hayyeh S, Nikolopoulou E, et al. Maternal cholestasis during pregnancy programs metabolic disease in offspring. *J Clin Invest*. 2013;123(7):3172-3181.
- Ridlon JM, Kang DJ, Hylemon PB, Bajaj JS. Bile acids and the gut microbiome. *Curr Opin Gastroenterol*. 2014;30(3):332-338.
- Oresic M, Simell S, Sysi-Aho M, et al. Dysregulation of lipid and amino acid metabolism precedes islet autoimmunity in children who later progress to type 1 diabetes. *J Exp Med*. 2008;205(13):2975-2984.
- la Marca G, Malvagia S, Toni S, Piccini B, Di Ciommo V, Bottazzo GF. Children who develop type 1 diabetes early in life show low levels of carnitine and amino acids at birth: does this finding shed light on the etiopathogenesis of the disease? *Nutr Diabetes*. 2013;3:e94.

26. Magnus P, Birke C, Vejrup K, et al. Cohort profile update: the Norwegian mother and child cohort study (MoBa). *Int J Epidemiol*. 2016;**45**(2):382-388.
27. Paltiel L, Haugan A, Skjerden T, et al. The biobank of the Norwegian Mother and Child Cohort Study—present status. *Nor J Epidemiol*. 2014;**24**(1-2):29-35.
28. Skrivarhaug T, Stene LC, Drivvoll AK, Strøm H, Joner G; Norwegian Childhood Diabetes Study Group. Incidence of type 1 diabetes in Norway among children aged 0-14 years between 1989 and 2012: has the incidence stopped rising? Results from the Norwegian Childhood Diabetes Registry. *Diabetologia*. 2014;**57**(1):57-62.
29. Størdal K, Mårild K, Tapia G, et al. Fetal and maternal genetic variants influencing neonatal vitamin D status. *J Clin Endocrinol Metab*. 2017;**102**(11):4072-4079.
30. Rønningen KS, Paltiel L, Meltzer HM, et al. The biobank of the Norwegian Mother and Child Cohort Study: a resource for the next 100 years. *Eur J Epidemiol*. 2006;**21**(8):619-625.
31. Tapia G, Suvitaival T, Ahonen L, et al. Data from: supplementary appendix to: prediction of type 1 diabetes at birth: cord blood metabolites versus genetic risk score in the MoBa cohort. Figshare 2021. Deposited May 10, 2021. doi:[10.6084/m9.figshare.14566476.v1](https://doi.org/10.6084/m9.figshare.14566476.v1).
32. Mårild K, Vistnes M, Tapia G, et al. Midpregnancy and cord blood immunologic biomarkers, HLA genotype, and pediatric celiac disease. *J Allergy Clin Immunol*. 2017;**139**(5):1696-1698.
33. Vistnes M, Tapia G, Mårild K, et al. Plasma immunological markers in pregnancy and cord blood: a possible link between macrophage chemo-attractants and risk of childhood type 1 diabetes. *Am J Reprod Immunol*. 2018;**79**(3):e12802.
34. Dilthey A, Leslie S, Moutsianas L, et al. Multi-population classical HLA type imputation. *Plos Comput Biol*. 2013;**9**(2):e1002877.
35. Aitken RJ, Mortimer GL, Gillespie KM. Type 1 diabetes high-risk HLA class II determination by polymerase chain reaction sequence-specific primers. *Methods Mol Biol*. 2016;**1433**:13-20.
36. Irgens LM. The Medical Birth Registry of Norway. Epidemiological research and surveillance throughout 30 years. *Acta Obstet Gynecol Scand*. 2000;**79**(6):435-439.
37. Chong J, Wishart DS, Xia J. Using metaboanalyst 4.0 for comprehensive and integrative metabolomics data analysis. *Curr Protoc Bioinformatics*. 2019;**68**(1):e86.
38. R Core Team. *R: a Language and Environment for Statistical Computing*. Vienna, Austria: R foundation for statistical computing; 2019.
39. Lamichhane S, Kempainen E, Trošt K, et al. Circulating metabolites in progression to islet autoimmunity and type 1 diabetes. *Diabetologia*. 2019;**62**(12):2287-2297.
40. Sen P, Dickens AM, López-Bascón MA, et al. Metabolic alterations in immune cells associate with progression to type 1 diabetes. *Diabetologia*. 2020;**63**(5):1017-1031.
41. Johnson RK, Vanderlinden L, DeFelice BC, et al.; TEDDY Study Group. Metabolite-related dietary patterns and the development of islet autoimmunity. *Sci Rep*. 2019;**9**(1):14819.
42. Li Q, Parikh H, Butterworth MD, et al.; TEDDY Study Group. Longitudinal metabolome-wide signals prior to the appearance of a first islet autoantibody in children participating in the TEDDY study. *Diabetes*. 2020;**69**(3):465-476.
43. Jørgenrud B, Stene LC, Tapia G, et al. Longitudinal plasma metabolic profiles, infant feeding, and islet autoimmunity in the MIDIA study. *Pediatr Diabetes*. 2017;**18**(2):111-119.
44. Pflueger M, Seppänen-Laakso T, Suoritti T, et al. Age- and islet autoimmunity-associated differences in amino acid and lipid metabolites in children at risk for type 1 diabetes. *Diabetes*. 2011;**60**(11):2740-2747.
45. Stanfill B, Reehl S, Bramer L, et al.; TEDDY Study Group. Extending classification algorithms to case-control studies. *Biomed Eng Comput Biol*. 2019;**10**:1179597219858954.
46. Webb-Robertson BM, Bramer LM, Stanfill BA, et al. Prediction of the development of islet autoantibodies through integration of environmental, genetic, and metabolic markers. *J Diabetes*. 2021;**13**(2):143-153.
47. Frohnert BI, Webb-Robertson BJ, Bramer LM, et al. Predictive modeling of type 1 diabetes stages using disparate data sources. *Diabetes*. 2020;**69**(2):238-248.
48. Storey JD. A direct approach to false discovery rates. *J R Stat Soc Ser B*. 2002;**64**(3):479-498.
49. Magnus MC, Tapia G, Olsen SF, et al. Parental smoking and risk of childhood-onset type 1 diabetes. *Epidemiology*. 2018;**29**(6):848-856.