

The Genetic Risk of Gestational Diabetes in South Asian women

Running title: Gestational Diabetes Genetic Risk in South Asian women

Amel Lamri (PhD)^{1,2}, Jayneel Limbachia (MSc)^{2,3}, Karleen M Schulze (MMath)², Dipika Desai (MSc)², Brian Kelly (PhD)⁴, Russell J de Souza (ScD)^{2,3}, Guillaume Paré (MD)^{2,3,5}, Deborah A Lawlor (PhD)⁶⁻⁸, John Wright (FRCP)⁴, Sonia S Anand (FRCPC)¹⁻³ for the Born in Bradford and START investigators

1. Department of Medicine, McMaster University and Hamilton Health Sciences, Hamilton, Ontario, Canada
2. Population Health Research Institute, McMaster University and Hamilton Health Sciences, Hamilton, Ontario, Canada
3. Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada
4. Bradford Institute for Health Research, Bradford Royal Infirmary, Bradford, UK
5. Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Ontario, Canada
6. Population Health Science, Bristol Medical School, University of Bristol, Bristol, UK.
7. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK.
8. Bristol NIHR Biomedical Research Centre, Bristol, UK

Corresponding Author: Dr. Sonia S Anand,
McMaster University, MDCL 3202,
1280 Main St W, Hamilton,
Ontario L8S 4K1, Canada.
Tel.: +1-905-525-9140, ext. 21523;
Fax: +1-905-528-2814.
e-mail: anands@mcmaster.ca

Word Count: 3629

Number of Tables and Figures: 4

ABBREVIATIONS

1KG	1000 Genomes
2hG	2h post-load Glucose
AF	Attributable Fraction
AUC	Area Under the Curve
AUCg	AUC glucose
BiB	Born in Bradford
CI	Confidence Interval
DIAGRAM	DIAbetes Genetics Replication and Meta-analysis
DNA	DeoxyriboNucleic Acid
FFQ	Food Frequency Questionnaire
FPG	Fasting Plasma Glucose
GCSE	General Certificate of Secondary Education
GDM	Gestational Diabetes Mellitus
GWAS	Genome-Wide Association Study
IADPSG	International Association of Diabetes and Pregnancy Study Groups
OR	Odds Ratio
PAF	Population Attributable Fraction
PRS	Polygenic Risk Score
SE	Standard Error
START	SouTh Asian biRth cohort
T2D	Type 2 Diabetes
UK	United Kingdom

ABSTRACT

South Asian women are at increased risk of developing gestational diabetes (GDM). Few studies have investigated the genetic contributions to GDM risk. We investigated the association of a type 2 diabetes (T2D) polygenic risk score (PRS), on its own, and with GDM risk factors, on GDM-related traits using data from two birth cohorts in which South Asian women were enrolled during pregnancy. 837 and 4,372 pregnant South Asian women from the South Asian Birth Cohort (START) and Born in Bradford (BiB) cohort studies underwent a 75-gram glucose tolerance test. PRSs were derived using GWAS results from an independent multi-ethnic study (~18% South Asians). Associations with fasting plasma glucose (FPG); 2h post-load glucose (2hG); area under the curve glucose; and GDM were tested using linear and logistic regressions. The population attributable fraction (PAF) of the PRS was calculated. Every 1 SD increase in the PRS was associated with a 0.085 mmol/L increase in FPG ([95%CI=0.07-0.10], $P=2.85 \times 10^{-20}$); 0.21 mmol/L increase in 2hG ([95%CI=0.16-0.26], $P=5.49 \times 10^{-16}$); and a 45% increase in the risk of GDM ([95%CI=32-60%], $P=2.27 \times 10^{-14}$), independent of parental history of diabetes and other GDM risk factors. PRS tertile 3 accounted for 12.5% of the population's GDM. No consistent interactions of the PRS with BMI, or diet quality were observed. A T2D PRS is strongly associated with multiple GDM-related traits in women of South Asian descent, and accounts for a substantial proportion of the PAF of GDM.

INTRODUCTION

Gestational diabetes mellitus (GDM) is defined as hyperglycaemia first diagnosed during pregnancy. This abnormal increase in blood glucose levels is associated with an increased risk of adverse health outcomes for both mother and their fetus/child during pregnancy, and later in life.(Farrar et al., 2016) It is estimated that 1% to >30% of live births are affected by GDM worldwide. This prevalence has been shown to vary widely depending on the participants ethnicity, countries/regions, and on the diagnostic criteria used.(Archambault, Arel, & Filion, 2014; McIntyre et al., 2019) South Asian women (whose ancestry derives from the Indian subcontinent) have a 2-fold increased odds of developing GDM, compared to white European women.(Anand et al., 2016; Cosson et al., 2014; Farrar et al., 2015; McIntyre et al., 2019) The reasons for this disproportionate risk have not been fully characterized.

Gestational diabetes is a complex disorder influenced by multiple genetic and environmental factors such as maternal age, ethnicity, obesity, poor diet quality, and family history of diabetes.(Anand et al., 2017; Hedderson, Darbinian, Quesenberry, & Ferrara, 2011; Solomon et al., 1997) Most genetic and environmental GDM risk factors are shared with type 2 diabetes ,(Sattar & Greer, 2002; Zhang & Ning, 2011) another condition which is thought to be very closely related to GDM. For example , women with GDM have a higher probability of having at least one parent with type 2 diabetes, compared to those with normal gestational glycemia.(Jang, Min, Lee, Cho, & Metzger, 1998) Furthermore, women with a GDM history have a 10-fold higher risk of subsequently being diagnosed with type 2 diabetes compared to those without a history of GDM.(Vounzoulaki et al., 2020) In terms of genetic architecture, both candidate gene and genome-

wide association studies (GWASs) demonstrated a considerable overlap between GDM and type 2 diabetes.(Hayes et al., 2013; Kwak et al., 2012; Pervjakova et al., 2021) Finally, type 2 diabetes polygenic risk scores (PRSs) have also been associated with GDM risk.(Lamri et al., 2020; Pervjakova et al., 2021)

It has been demonstrated that environmental exposures such as diet and/or physical activity may modulate the effect of type 2 diabetes loci (such as *TCF7L2*, *PPARG* and *CDKALI*) on the risk of type 2 diabetes.(Dietrich et al., 2019) Nevertheless, only a handful of studies have investigated genetic×environmental interactions on GDM,(Chen et al., 2019; Grotenfelt et al., 2016; Popova et al., 2017) and to date, no study has tested the interaction between a genome-wide PRS with other GDM risk factors, on the risk of GDM.

The aims of this investigation were to: *i*) test the association of a type 2 diabetes PRS, generated from an external multi-ethnic GWAS (~18% South Asians), with GDM and related traits (fasting plasma glucose (FPG), 2h post-load glucose (2hG) and area under the curve glucose (AUCg) levels) in pregnant South Asian women from the South Asian birth cohort (START) and the Born in Bradford (BiB) studies; *ii*) To estimate the population attributable fraction (PAF) of the PRS on GDM, and *iii*) To determine whether the effect of the PRS is modulated by other GDM risk factors including age, BMI, diet quality, birth country, education, and parity.

METHODS

Study design and participants

START is a prospective cohort study designed to evaluate the environmental and genetic determinants of cardio-metabolic traits among South Asian women and their offspring living in Canada.(Anand et al., 2013) In brief, 1,012 South Asian pregnant women, aged between 18 and 40 years old, were recruited during their second trimester of pregnancy from the Peel Region (Ontario, Canada) through physician referrals between 2011 and 2015. All START participants provided informed consent, and the study was approved by local ethics committees (Hamilton Integrated Research Ethics Board [ID:10-640], William Osler Health System [ID:11-0001], and Trillium Health Partners [RCC:11-018, ID:492]).

BiB is a prospective, longitudinal family cohort study designed to investigate the causes of illness, and develop interventions to improve health in a deprived multi-ethnic population in Bradford, England, UK.(Wright et al., 2013) Between 2007 and 2011, 12,453 women of various ethnic backgrounds (~46% South Asian origin) were recruited between their 24th and 28th week of pregnancy. Detailed information on socio-economic characteristics, ethnicity, family history, environmental and physical risk factors has been collected.(Farrar et al., 2015; Wright et al., 2013) Ethical approval for all aspects of the research was granted by Bradford Research Ethics Committee [Ref 07/H1302/112].

Measurements and questionnaires:

START: A detailed description of the maternal measurements has been published previously.(Anand et al., 2017) Briefly, weight and height were measured using standard procedures, and information about pre-pregnancy weight, family and personal medical history was collected using questionnaires. Parental history of diabetes was derived from baseline questionnaires and categorized as neither parents had a history of diabetes or either one or both

parents did. Birth country, number of years spent in Canada, and education-related variables were self-reported. Participants' highest level of education was coded as a five-category ordinal variable as: 1) less than high school; 2) high school completed; 3) Diploma or certificate from trade, technical or vocational school; 4) Bachelor's or undergraduate degree, or teacher's college; 5) Master's, Doctorate or professional degree. A binary "born in South Asia" variable was categorized as participants born in South Asia (India, Pakistan, Sri Lanka or Bangladesh versus participants were born in any other country. A validated ethnic-specific food frequency questionnaire (FFQ) was used to collect dietary information.(Kelemen et al., 2003) To calculate diet quality, 1 point was given for consuming \geq the study population median of: 1) green leafy vegetables, 2) raw vegetables, 3) other cooked vegetables, and 4) fruits, and less than the population median of 5) fried foods/fast food/ or snacks, and 6) meat/poultry. The continuous scores (ranging from 0 to 6 points) was categorized into a binary variable: low diet quality vs. medium or high quality] before analysis.(Anand et al., 2017)

BiB: Maternal height was measured during the recruitment visit (24-28th weeks of pregnancy) using standard procedures. In the absence of pre-pregnancy weight data, weight from the first antenatal clinic visit (average 12 weeks of pregnancy) was used to calculate BMI. Ethnicity of participants and years spent in UK were self-reported at recruitment through an interview administered questionnaire; missing ethnicity data was backfilled from primary care data when available. The South Asian ethnicity of all participants included in this analysis was validated using genetic data. Parental history of diabetes and "born in South Asia" variables were derived from the baseline questionnaire data and coded as in START. Since only a very small proportion of BiB's participants completed an FFQ that included information about fruits and vegetables intake, the diet quality score could not be derived in BiB. Data regarding the participant's highest

educational qualification was equalized (using UK standards) and recoded into the following categories: 1) less than 5 General Certificate of Secondary Education (GCSE) equivalent; 2) 5 GCSE equivalent; 3) A-level equivalent; 4) higher than A-level. Data for unclassifiable foreign degrees was considered as missing.

Outcomes: Study participants without prior type 2 diabetes were invited to undertake a 75-gram oral glucose tolerance test (OGTT) in both START and BiB, and FPG, and 2hG levels were measured (1h post-load glucose was measured in START only). AUCg was calculated using the FPG and 2hG glucose levels in BiB, and using the FPG, 1h post-load glucose, and 2hG levels in START.(Anand et al., 2017) Given the difference in the number of data points included in the calculation of AUC between the two studies and the skewness of the distributions, values were log-transformed, winsorized and standardized in each study before analysis. Gestational diabetes status of women without pre-existing type 2 diabetes was primarily defined based on OGTT results in both studies using the International Association of Diabetes and Pregnancy Study Group (IADPSG) GDM criteria (FPG ≥ 5.1 mmol/L or higher, or a 1hG ≥ 10.8 or a 2hG ≥ 8.5 mmol/L or higher)(International Association of et al., 2010). Our secondary outcome was GDM using BiB's South Asian specific definition (FPG of 5.2 mmol/L or higher, or a 2hG of 7.2 mmol/L or higher) (Farrar et al., 2015), which will be referred to as the South Asian-specific definition hereafter. Self-reported GDM status or data from the birth chart was used to determine GDM's status if OGTT measures were unavailable (N=65 and 31 in START and BiB respectively). Women with pre-existing diabetes at baseline were not included in this analysis. Pre-pregnancy diabetes status was determined using maternal self-reported data (about diabetes diagnosis, diabetes medication and/or insulin intake prior to pregnancy) in START. In BiB, information on pre-pregnancy diabetes was backfilled from electronic medical records.

In order to keep a single pregnancy (and a single GDM status) per mother in BiB, only pregnancies with no missing data for GDM were included. For mothers with available data at multiple pregnancies at this stage, pregnancies with no missing data across all covariates (age, BMI, family history, birth country, parity, education level) were prioritized. Next, only pregnancies with the least amount of missing data across all covariates were kept. The following two additional filtering approaches were then applied for mothers with multiple pregnancies remaining: i) if GDM was not diagnosed at any of the pregnancies, phenotype data at the latest available time point was kept (ie. keep older GDM controls). ii) if GDM was diagnosed during any of the pregnancies included in the study, the earliest time point where GDM was diagnosed was kept (ie. keep younger GDM cases).

DNA extraction, Genotyping, Imputation and Filtering:

START: DNA was extracted and genotyped for 867 mothers using the Illumina Human CoreExome-24 and Infinium CoreExome-24 arrays (Illumina, San-Diego, CA, USA). 837 samples passed standard quality control procedures.(Anderson et al., 2010) Genotypes were phased and imputed using SHAPEIT v2.12(Delaneau, Marchini, Genomes Project, & Genomes Project, 2014), and IMPUTE v2.3.2(Howie, Donnelly, & Marchini, 2009) respectively using the 1000 Genomes (phase 3) data as a reference panel.(Consortium et al., 2015) Variants with an info score <0.7 were removed from analysis. In total, 837 START participants with both genotypes and available GDM status, FPG, 1h- and/or 2hG levels were included in the analysis (Figure S1).

Born in Bradford: DNA was extracted and genotyped for 16,267 and 3,663 BiB participants using the Illumina HumanCoreExome (12v1.0, 12v1.1 or 24v1.0) and InfiniumGlobal Screening Array

(24v2.0) arrays respectively (Illumina, San-Diego, CA, USA). 4,372 South Asian mothers passed genotyping quality controls, had GDM status, FPG, and/or 2hG levels available, and were included in our analysis (Figure S1).

Deriving the PRS: Given the absence of publicly available South Asian-specific T2D or GDM GWAS data at the time of the analysis, weights were derived from the DIAGRAM's 2014 multi-ethnic T2D GWAS meta-analysis, which included over 18% of South Asians (~63% European and 19% other ethnic backgrounds).(Mahajan et al., 2014) A grid search approach was used to identify the optimal parameters (17 P-values tested, ranging from 5×10^{-8} to 1 with 0.1 increase; 4 heritability values tested: 0.023; 0.06; 0.08; 0.12). START and BiB genotypes were pooled. 70% of the samples' data were used for training and 30% for validation (random sampling stratified by study) in order to minimize the impact of population stratification. The PRS was derived using LDpred2.(Prive, Arbel, & Vilhjalmsson, 2020) The best PRS (i.e. that maximized the AUC) was characterized by a P value ≤ 0.0014 and an $h^2=0.08$ ($N_{SNVs}=6,492$). The PRS was standardized (mean=0, standard deviation=1) in both studies before analysis.

Principal component analysis of genetic data

A principal component analysis (PCA) was performed using the PC-Air function from the GENESIS R package (v2.20.0).(Conomos, Miller, & Thornton, 2015) Kinship matrices (required to derive PCs with PC-Air) were derived using KING (v2.2.5).(Manichaikul et al., 2010)

Statistical Analysis

The statistical analysis was conducted using R (v3.6.3).(R core Team, 2016) Linear regression models were used to test the association between the PRS and FPG, 2hG and AUCg. PRS and

GDM associations were tested using logistic regression. Both univariate and multivariate models were constructed with adjustment for GDM risk factors (age, BMI, parity, birth in South Asia (yes vs. no), education level, and diet quality (in START only) and the first 5 PCs (in order to minimize the effect of population stratification). Interactions between the PRS and each risk factors was also tested. The estimated population attributable fractions (PAFs) and their corresponding standard errors were calculated using the AF R package (v.0.1.5). To this end, continuous variables were recoded into categorical variables: age was divided in two categories [(29-31, 32-43) vs. 19-28]; BMI was stratified into a two categories variable using South Asian obesity cutoff points suggested by Gray et al. (Gray et al., 2011) (<23 vs. ≥23); The PRS was divided into two categories (tertiles 1+2 vs tertile 3); Parity was divided into two categories (primiparity vs. 1 pregnancy or more); Education level variables were divided in two categories (completed high school or lower vs. higher degree, diploma or certificate in START; and A-level equivalent or lower vs. higher than A-level in BiB).

RESULTS

The proportion of women classified with GDM using the IADPSG criteria was 25% and 11.2% in START and BiB respectively, which was lower than the proportion using the South Asian-specific definition of 36.2% and 22.9% respectively. Notably the proportion of women with GDM was higher in START compared to BiB irrespective of the classification method used.

The proportion of women of Indian origin in START and BiB was 71.8% and 5.1%, while the proportion of Pakistani women was 23.4% and 94.3% respectively. The proportion of participants born in the Indian sub-continent was higher in START (88.6%) than in BiB (55.6%), and the

average number of years spent in Canada or the UK among these participants was lower in START compared to BiB (6.6 vs. 9.7 years respectively). The proportions of primiparous women (40.9% vs 31.7%) and women with 1 prior pregnancy (42.4% vs 26.9%) were higher in START than in BiB. Conversely, participants with 2 or more prior pregnancies were more frequent in BiB than START (41.4% vs. 16.6% respectively). The proportion of vegetarian participants was higher in START than in BiB (36.4% vs 1.3%). Finally, the proportion of participants with a post-secondary degree/diploma or higher was greater in START than BiB (84.0% vs 29.0%).

Table 1 shows the baseline characteristics of the South Asian women from the START and BiB stratified by GDM case vs non GDM (IADPSG criteria). As expected, women with GDM had a higher mean fasting, 2hG and AUCg levels than non-GDM participants. Participants with GDM were older, had a higher BMI, and were more likely to report a family history of diabetes compared to women without GDM, in both studies. The overall diet quality was lower in participants with GDM compared to non-GDM participants in START (data not available in BiB). Of note, the average difference in BMI between GDM cases and controls was higher in BiB than in START (2.8 and 1.8 respectively) (Table 1).

The standardized PRS ranged between -3.23 and 3.12 in START as compared to -3.51 and 4.16 in BiB. The full list of genetic variants included in the PRS as well as their characteristics are shown in Table S1. Women with GDM had a higher mean PRS compared to women without GDM. Similarly, women with GDM were more likely to have PRS categorized in tertile 2 or 3, compared to tertile 1 (Table 1).

Genetic risk and GDM related traits in univariate models

The continuous PRS was associated with FPG, 2hG, and AUCg in START and BiB in univariate models. Every 1 SD increase in the PRS was associated with a 0.09 mmol/L increase in FPG [95%CI=0.07-0.10], 0.23 mmol/L increase in 2hG [95%CI=0.18-0.28], and a 0.17 unit increase in AUCg z-scores [0.14 - 0.20] in the meta-analysed results (Table S2).

The PRS was also associated with the risk of GDM_{IADPSG} in univariate models whereby a 1 SD increase in PRS was associated with a 47% increase in risk of GDM after meta-analysis [95%CI=35-60%]. A similar association is observed using the South Asian-specific definition of GDM, with moderate between-study heterogeneity observed (Table S2).

Overall, the risk of GDM_{IADPSG} increased progressively comparing tertile 2 of the PRS to tertile 1, and tertile 3 to tertile 1 (43% and 230% respectively, Table S2). Higher PRS categories were also associated with higher FPG, 2hG and AUCg levels (Table S2).

Multivariable models of GDM risk factors and GDM related traits.

The continuous PRS was strongly and independently associated with FPG, 2hG and AUCg levels in a multivariable model adjusted for age, BMI, parity, parental history of diabetes, region of birth (South Asia vs. other), education level, and diet quality (available in START only), and the first 5 PCs (Table 2). For example, every 1 SD increase in the PRS was associated with a 0.08 mmol/L increase in FPG, and 0.21 mmol/L increase in 2hG levels (Table 2). The continuous PRS was also associated with a higher risk of GDM in a model with similar adjustments whereby every 1 SD increase in the PRS was associated with a 45% increase in the risk of GDM_{IADPSG} (Table 2). Association results for GDM using the South Asian-specific criteria are shown in Table S3

When testing tertiles of PRS with similar covariates, our results show that participants in the 2nd and 3rd PRS tertiles have a 37% and 19% increase in the risk of GDM_{IADPSG} compared to participants in tertile 1 respectively (Table S4). Higher PRS tertiles were also associated with higher FPG, 2hG and AUCg levels (Table S4). The effect sizes associated with tertiles 2 were higher in START than BiB across multiple GDM related traits (2hG, AUCg and GDM, Table S4).

Population attributable fraction

In a model adjusted for maternal age, BMI, education, birth in South Asia (yes/no), parental history of diabetes, and diet quality (in START only), the PRS tertile 3 accounted for 12.5% of the population's total GDM_{IADPSG} cases overall, and was higher in START than in BiB (Table 3). The combined effect of PRS and parental history of diabetes on GDM accounted for ~21.7% of the population's GDM cases in the two studies combined (Table 3).

Interactions between the PRS and GDM risk factors on GDM

No consistent interactions were observed between the PRS and maternal age; parity; or education level modulating FPG, 2hG, AUCg, or GDM_{IADPSG} in START or BiB (Table 4 and Table S5).

Some nominally significant interactions modulating the continuous trait of FPG were observed in START which were not confirmed in BiB. These included the PRS with BMI and the PRS×birth in South Asia (yes/no) interactions ($P_{\text{interaction}}=0.01$ and 0.04 respectively), yet non-significant in BiB ($P_{\text{interaction}}=0.05$ and 0.07 respectively), with different effect sizes between the two studies (Table S6), resulting in non-significant meta-analysis of these effects ($P_{\text{interaction}}=0.45$ and 0.26 respectively). Lastly, a PRS×diet quality interaction on FPG was detected in START ($P_{\text{interaction}}=0.002$, Table 4) whereby the effect of the PRS appeared to be stronger in participants

with a low diet quality (Beta=0.17 [95%CI=0.10-0.24]) than in participants with a medium or high diet quality (Beta=0.05 [95%CI=0.00-0.09]) (Table S6 and Figure S2). The overall diet quality score was not available BiB, hence this interaction could not be tested for replication.

DISCUSSION

We demonstrate that a type 2 diabetes polygenic risk score, based on an independent and multi-ethnic GWAS meta-analysis (with 18% South Asian participants), is strongly associated with GDM and related glucose traits among South Asian pregnant women settled in Canada or the UK. This association is independent of other known GDM risk factors, including maternal age, BMI, parental history of diabetes, and maternal birth country. The PRS highest tertile accounted for 13.8 % of the PAF of GDM. Consistent with a recent trans-ethnicity GWAS of GDM, and these results support the hypothesis that GDM and type 2 diabetes are part of the same underlying pathology.(Pervjakova et al., 2021)

Family history of type 2 diabetes is often used as a surrogate marker of the genetic risk of type 2 diabetes. Our results show that the addition of the PRS to the multivariate models does not nullify the impact of parental history on GDM and *vice versa*. This suggests that the PRS and family history of diabetes both partially convey independent information. This partial independence could be explained by the fact that the PRS does not entirely capture the genetic association signals with GDM. On the other hand, family history reflects not only genetic similarity, but also shared non-genetic lifestyle factors.

Overall, we found no robust evidence for modulation of the PRS's effect on GDM-related traits by other GDM risk factors. While marginal PRS×BMI and PRS×South Asia born interactions on FPG were observed, these did not replicate, both in terms of statistical significance and effect sizes, suggesting that these environmental and genetic factors may act independently. Furthermore, these interactions would not pass multiple testing corrections if applied. A potentially stronger PRS×diet quality interaction modulating FPG was observed in START. However, since it was not possible to replicate this interaction in BiB, future investigations are required in order to validate this observation. If confirmed, this interaction may help identify a subpopulation who will benefit the most from a targeted diet intervention for the prevention of GDM.

The overall clinical implications of our findings should be carefully considered. At present, the use of laboratory-derived genetic information in the clinical setting remains expensive and is not implemented for complex diseases like GDM. Given the enhanced predictive value of our genome-wide PRS, future evaluations of whether the knowledge of one's genetic risk improves adherence to lifestyle recommendations and contributes to reducing gestational hyperglycemia would be of great interest.

Our study has been considerably strengthened by the use of a PRS optimized for a large population of South Asians from two independent cohorts, as well as by the fact that GDM status was determined using objective OGTT measures. Nevertheless, there are some limitations to our analysis that should be considered: *i*) The weights attributed to the genetic variants included in the PRS are derived from a type 2 diabetes study. Overall, evidence points to a strong correlation between top variants from type 2 diabetes and GDM GWASs. However, variants at some common loci (eg. *MTNR1B*) might have significantly different effect size depending on the phenotype

studied.(Pervjakova et al., 2021) In addition, variants in at least one locus (*HKDC1*) have been strongly associated to GDM but not type 2 diabetes.(Pervjakova et al., 2021) More GDM-specific loci, or loci with a different magnitude of effect between GDM and type 2 diabetes might be identified from future, larger studies. These observations suggest that future PRSs based on a GDM GWAS may have a higher predictive value and more power to detect gene×environment interactions. *ii*) Second, some differences in measurements exist between START and BiB studies, including the timing of weight measurements, and the number of data points included in the calculation of AUCg. However, since data was standardized in both studies, we do not expect that AUCg measurements differences had a major impact on the results. *iii*) Finally, the comparison of genetic data between START and BiB revealed the existence of genetic heterogeneity, both between and within the samples of these two cohorts (Figure S3). It is our assumption that these differences can be explained by the difference of sample size (START being smaller than BiB), as well as by historical differences in migration patterns from South Asia to Canada and the UK. For example, most START participants were first generation migrants from India, whereas the majority of South Asians in BiB are descendants of Pakistani migrants who settled in the UK for several generations. In order to account for this genetic heterogeneity, we derived a new type 2 diabetes PRS that combined samples from the two studies. This PRS should be more generalizable to other South Asian studies. Another measure implemented to reduce the effect of population stratification was the adjustment for the PC axes in our analysis. Given the absence of heterogeneity in our FPG, 2hG, or GDM_{IADPSG} PC adjusted models, we consider that population stratification effects have been accounted for.

296 *Conclusion:* A type 2 diabetes derived PRS is strongly associated with the risk of GDM in pregnant
297 women of South Asian descent, independent of parental history of diabetes, and other GDM risk
298 factors.

Acknowledgements

Research projects in START and Born in Bradford are only possible because of the enthusiasm and commitment of the parents and children involved in these two studies. We are grateful to all the participants, teachers, school staff, health professionals and researchers and other contributors who have made these studies happen.

Funding

Studies: The South Asian Birth Cohort (**START**) study data were collected as part of a program funded by the Indian Council of Medical Research in Canada and by the Canadian Institutes of Health Research (grant INC-109205), and the Heart and Stroke Foundation (grant NA7283) with founding principal investigators: Sonia S Anand, Anil Vasudevan, Milan Gupta, Katherine Morrison, Anura Kurpad, Koon K Teo, and Krishnamachari Srinivasan.

The Born in Bradford (**BiB**) The Born in Bradford cohort is funded by the National Institute for Health Research Collaboration for Applied Health Research and Care (NIHR CLAHRC) and the Programme Grants for Applied Research funding scheme (RP-PG-0407-10044). The study also receives funding from the Wellcome Trust (WT101597MA), a joint grant from the UK Medical Research Council (MRC) and Economic and Social Science Research Council (ESRC) (MR/N024397/1) and the British Heart Foundation (CS/16/4/32482). DNA extraction was funded by the UK Medical Research Council via the Integrative Epidemiology Unit (MRC IEU; MC_UU_12013/5) and genotyping via the MRC IEU and a National Institute of Health Research Senior Investigator Award to D.A.L. (NF-0616-10102).

Authors: Research associate (**A.L.**) and graduate student (**J.L.**) costs were covered by two Canadian Institutes of Health Research Grants [Project grant number: 298104, Foundation Scheme

grant number: FDN-143255, Study grant numbers: INC 109205, NA 7283] awarded to S.S.A; **D.A.L.**'s contribution to this study is supported by the Bristol NIHR Biomedical Research Centre, the UK Medical Research Council (MC_UU_00011/6) and the British Heart Foundation (CH/F/20/90003). **S.S.A.** is supported by a Tier 1 Canada Research Chair in Ethnic Diversity and Cardiovascular Disease, and a Heart and Stroke Foundation/Michael G. DeGroote Chair in Population Health Research at McMaster University.

Competing/Duality of interests

D.A.L. has received support from Medtronic Ltd and Roche Diagnostics for research unrelated to that presented here. No financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work. Other than those declared by **D.A.L.** above, no authors have any conflict of interest.

Author contributions

A.L. derived the PRS, performed the association and interaction analysis, drafted and revised the manuscript. **J.L.** performed the association tests and PAR analysis in START, drafted and revised the manuscript. **K.S.** reviewed and supervised the statistical analysis and commented on the paper. **D.D.** is the study coordinator for the START birth cohort and provided comments on the manuscript. **B.K.** directed BiB data acquisition, provided comments on statistical analysis and reviewed the manuscript, **R.J.D.** provided comments on statistical analysis and reviewed manuscript. **G.P.** performed the genotyping laboratory analysis and provided comments on the manuscript. **D.A.L.** provided comments on the analysis and reviewed the manuscript. J.R. the BiB

study's principal investigator, he provided comments on statistical analysis and reviewed the manuscript. S.S.A. (guarantor) oversaw the development, analysis, and writing of this manuscript.

Data and code availability

Data from START is not publicly available, since the study is bound by consent which indicates the data will not be used by an outside group. Requests for collaboration or replication will be considered for research purposes only (no commercial use allowed, as per the study's informed consent). Requests should be addressed to the study's principal investigator (Sonia Anand, anands@mcmaster.ca) via a form which will be provided upon request by emailing natcampb@mcmaster.ca. The request will be evaluated by PIs and co-investigators, and projects deemed of scientific interest will be further evaluated/validated by local REB chair. Born in Bradford data are available for research purposes only by sending an expression of interest form downloadable from https://borninbradford.nhs.uk/wp-content/uploads/BiB_EoI_v3.1_10.05.21.doct to borninbradford@bthft.nhs.uk. The proposal will be reviewed by BiB's executive team. If the request is approved, the requester will be asked to sign a [Data Sharing Contract](#) and a [Data Sharing Agreement](#). Full details on how to access data and forms can be found here <https://borninbradford.nhs.uk/research/how-to-access-data/>. The code used to analyze the data is available at https://github.com/AmelLamri/Paper_T2dPrsGdm_StartBiB. All Sharable processed versions of the datasets used in the manuscript are made available as supplementary material or at https://github.com/AmelLamri/Paper_T2dPrsGdm_StartBiB.

REFERENCES

- Anand, S. S., Gupta, M., Teo, K. K., Schulze, K. M., Desai, D., Abdalla, N., . . . South Asian Birth Cohort - Canada, I. (2017). Causes and consequences of gestational diabetes in South Asians living in Canada: results from a prospective cohort study. *CMAJ Open*, 5(3), E604-E611. doi:10.9778/cmajo.20170027
- Anand, S. S., Gupta, M. K., Schulze, K. M., Desai, D., Abdalla, N., Wahi, G., . . . Teo, K. K. (2016). What accounts for ethnic differences in newborn skinfold thickness comparing South Asians and White Caucasians? Findings from the START and FAMILY Birth Cohorts. *Int J Obes (Lond)*, 40(2), 239-244. doi:10.1038/ijo.2015.171
- Anand, S. S., Vasudevan, A., Gupta, M., Morrison, K., Kurpad, A., Teo, K. K., . . . Investigators, S. C. S. (2013). Rationale and design of South Asian Birth Cohort (START): a Canada-India collaborative study. *BMC Public Health*, 13, 79. doi:10.1186/1471-2458-13-79
- Anderson, C. A., Pettersson, F. H., Clarke, G. M., Cardon, L. R., Morris, A. P., & Zondervan, K. T. (2010). Data quality control in genetic case-control association studies. *Nat Protoc*, 5(9), 1564-1573. doi:10.1038/nprot.2010.116
- Archambault, C., Arel, R., & Filion, K. B. (2014). Gestational diabetes and risk of cardiovascular disease: a scoping review. *Open Med*, 8(1), e1-9.
- Chen, Y., Liu, H., Wang, L., Zhou, T., Liang, Z., Li, W., . . . Qi, L. (2019). Lifestyle intervention modifies the effect of the MC4R genotype on changes in insulin resistance among women with prior gestational diabetes: Tianjin Gestational Diabetes Mellitus Prevention Program. *Am J Clin Nutr*. doi:10.1093/ajcn/nqz121
- Conomos, M. P., Miller, M. B., & Thornton, T. A. (2015). Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. *Genet Epidemiol*, 39(4), 276-293. doi:10.1002/gepi.21896
- Consortium, T. G. P., Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., . . . Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*, 526(7571), 68-74. doi:10.1038/nature15393
- Cosson, E., Cussac-Pillegand, C., Benbara, A., Pharisien, I., Jaber, Y., Banu, I., . . . Carbillon, L. (2014). The diagnostic and prognostic performance of a selective screening strategy for gestational diabetes mellitus according to ethnicity in Europe. *J Clin Endocrinol Metab*, 99(3), 996-1005. doi:10.1210/jc.2013-3383
- Delaneau, O., Marchini, J., Genomes Project, C., & Genomes Project, C. (2014). Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. *Nat Commun*, 5, 3934. doi:10.1038/ncomms4934

- Dietrich, S., Jacobs, S., Zheng, J. S., Meidtner, K., Schwingshackl, L., & Schulze, M. B. (2019). Gene-lifestyle interaction on risk of type 2 diabetes: A systematic review. *Obes Rev*. doi:10.1111/obr.12921
- Farrar, D., Fairley, L., Santorelli, G., Tuffnell, D., Sheldon, T. A., Wright, J., . . . Lawlor, D. A. (2015). Association between hyperglycaemia and adverse perinatal outcomes in south Asian and white British women: analysis of data from the Born in Bradford cohort. *Lancet Diabetes Endocrinol*, 3(10), 795-804. doi:10.1016/S2213-8587(15)00255-7
- Farrar, D., Simmonds, M., Bryant, M., Sheldon, T. A., Tuffnell, D., Golder, S., . . . Lawlor, D. A. (2016). Hyperglycaemia and risk of adverse perinatal outcomes: systematic review and meta-analysis. *BMJ*, 354, i4694. doi:10.1136/bmj.i4694
- Gray, L. J., Yates, T., Davies, M. J., Brady, E., Webb, D. R., Sattar, N., & Khunti, K. (2011). Defining obesity cut-off points for migrant South Asians. *PLoS One*, 6(10), e26464. doi:10.1371/journal.pone.0026464
- Grotenfelt, N. E., Wasenius, N. S., Rono, K., Laivuori, H., Stach-Lempinen, B., Orho-Melander, M., . . . Eriksson, J. G. (2016). Interaction between rs10830963 polymorphism in MTNR1B and lifestyle intervention on occurrence of gestational diabetes. *Diabetologia*, 59(8), 1655-1658. doi:10.1007/s00125-016-3989-1
- Hayes, M. G., Urbanek, M., Hivert, M. F., Armstrong, L. L., Morrison, J., Guo, C., . . . Group, H. S. C. R. (2013). Identification of HKDC1 and BACE2 as genes influencing glycemic traits during pregnancy through genome-wide association studies. *Diabetes*, 62(9), 3282-3291. doi:10.2337/db12-1692
- Hedderson, M. M., Darbinian, J. A., Quesenberry, C. P., & Ferrara, A. (2011). Pregravid cardiometabolic risk profile and risk for gestational diabetes mellitus. *Am J Obstet Gynecol*, 205(1), 55 e51-57. doi:10.1016/j.ajog.2011.03.037
- Howie, B. N., Donnelly, P., & Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*, 5(6), e1000529. doi:10.1371/journal.pgen.1000529
- International Association of, D., Pregnancy Study Groups Consensus, P., Metzger, B. E., Gabbe, S. G., Persson, B., Buchanan, T. A., . . . Schmidt, M. I. (2010). International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. *Diabetes Care*, 33(3), 676-682. doi:10.2337/dc09-1848
- Jang, H. C., Min, H. K., Lee, H. K., Cho, N. H., & Metzger, B. E. (1998). Short stature in Korean women: a contribution to the multifactorial predisposition to gestational diabetes mellitus. *Diabetologia*, 41(7), 778-783. doi:10.1007/s001250050987
- Kelemen, L. E., Anand, S. S., Vuksan, V., Yi, Q., Teo, K. K., Devanesen, S., . . . Investigators, S. (2003). Development and evaluation of cultural food frequency questionnaires for South Asians, Chinese, and Europeans in North America. *J Am Diet Assoc*, 103(9), 1178-1184. doi:10.1053/jada.2003.50578

- Kwak, S. H., Kim, S. H., Cho, Y. M., Go, M. J., Cho, Y. S., Choi, S. H., . . . Park, K. S. (2012). A genome-wide association study of gestational diabetes mellitus in Korean women. *Diabetes*, 61(2), 531-541. doi:10.2337/db11-1034
- Lamri, A., Mao, S., Desai, D., Gupta, M., Pare, G., & Anand, S. S. (2020). Fine-tuning of Genome-Wide Polygenic Risk Scores and Prediction of Gestational Diabetes in South Asian Women. *Sci Rep*, 10(1), 8941. doi:10.1038/s41598-020-65360-y
- Mahajan, A., Go, M. J., Zhang, W., Below, J. E., Gaulton, K. J., Ferreira, T., . . . Type 2 Diabetes Genetic Exploration by Nex-generation sequencing in muylti-Ethnic Samples, C. (2014). Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet*, 46(3), 234-244. doi:10.1038/ng.2897
- Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W. M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics*, 26(22), 2867-2873. doi:10.1093/bioinformatics/btq559
- McIntyre, H. D., Catalano, P., Zhang, C., Desoye, G., Mathiesen, E. R., & Damm, P. (2019). Gestational diabetes mellitus. *Nat Rev Dis Primers*, 5(1), 47. doi:10.1038/s41572-019-0098-8
- Pervjakova, N., Moen, G.-H., Borges, M.-C., Ferreira, T., Cook, J. P., Allard, C., . . . Mägi, R. (2021). Trans-ancestry genome-wide association study of gestational diabetes mellitus highlights genetic links with type 2 diabetes. *medRxiv*. doi:10.1101/2021.10.11.21264235
- Popova, P. V., Klyushina, A. A., Vasilyeva, L. B., Tkachuk, A. S., Bolotko, Y. A., Gerasimov, A. S., . . . Grineva, E. N. (2017). Effect of gene-lifestyle interaction on gestational diabetes risk. *Oncotarget*, 8(67), 112024-112035. doi:10.18632/oncotarget.22999
- Prive, F., Arbel, J., & Vilhjalmsen, B. J. (2020). LDpred2: better, faster, stronger. *Bioinformatics*. doi:10.1093/bioinformatics/btaa1029
- R core Team. (2016). R: A language and environment for statistical computing (Version 3.3). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Sattar, N., & Greer, I. A. (2002). Pregnancy complications and maternal cardiovascular risk: opportunities for intervention and screening? *BMJ*, 325(7356), 157-160. doi:10.1136/bmj.325.7356.157
- Solomon, C. G., Willett, W. C., Carey, V. J., Rich-Edwards, J., Hunter, D. J., Colditz, G. A., . . . Manson, J. E. (1997). A prospective study of pregravid determinants of gestational diabetes mellitus. *JAMA*, 278(13), 1078-1083.
- Vounzoulaki, E., Khunti, K., Abner, S. C., Tan, B. K., Davies, M. J., & Gillies, C. L. (2020). Progression to type 2 diabetes in women with a known history of gestational diabetes: systematic review and meta-analysis. *BMJ*, 369, m1361. doi:10.1136/bmj.m1361

- Wright, J., Small, N., Raynor, P., Tuffnell, D., Bhopal, R., Cameron, N., . . . Born in Bradford Scientific Collaborators, G. (2013). Cohort Profile: the Born in Bradford multi-ethnic family cohort study. *Int J Epidemiol*, 42(4), 978-991. doi:10.1093/ije/dys112
- Zhang, C., & Ning, Y. (2011). Effect of dietary and lifestyle factors on the risk of gestational diabetes: review of epidemiologic evidence. *Am J Clin Nutr*, 94(6 Suppl), 1975S-1979S. doi:10.3945/ajcn.110.001032

URLs

1000 Genomes study

<https://www.internationalgenome.org/>

Born in Bradford Study

<https://borninbradford.nhs.uk/>

DIAGRAM consortium

<https://www.diagram-consortium.org/>

GCTA: a tool for Genome-wide Complex Trait Analysis

<https://cnsgenomics.com/software/gcta>

KING: Relationship Inference

<https://people.virginia.edu/~wc9c/KING/manual.html>

PC-AiR: Principal Components Analysis in Related Samples

<https://rdrr.io/bioc/GENESIS/man/pcair.html>

PLINK

<https://www.cog-genomics.org/plink/>

R Project

<https://cran.r-project.org/>

TABLES

Table 1: Characteristics of START and BiB study participants included in the analysis.

	START			BiB		
	No GDM	GDM	P-value	No GDM	GDM	P-value
N (%)	759 (75)	253 (25)	-	3809 (88.8)	481 (11.2)	-
Age, years	29.8 (3.8)	31.6 (4)	5.55×10^{-10}	27.7 (5)	30.5 (5.4)	1.40×10^{-22}
Height, cm	162.5 (6.27)	161.13 (6.01)	0.002	159.9 (5.69)	158.3 (5.66)	6.19×10^{-08}
Weight, kg ^a	61.7 (11.7)	65.6 (12.9)	2.00×10^{-05}	64.8 (14.1)	71.1 (15.1)	2.22×10^{-16}
BMI, kg/m ² ^b	23.4 (4.3)	25.3 (4.9)	4.93×10^{-08}	25.4 (5.2)	28.4 (5.8)	5.89×10^{-23}
Parity, n (%)						
0	328 (44.3%)	78 (31.1%)		1189 (32.2%)	129 (27.4%)	
1	299 (40.4%)	122 (48.6%)	0.001	1026 (27.8%)	94 (20%)	5.03×10^{-07}
2 or more	114 (15.4%)	51 (20.3%)		1473 (39.9%)	248 (52.7%)	
Post-secondary education, n (%)	641 (84.6%)	208 (82.2%)	0.43	952 (29.4%)	110 (26.3%)	0.22
Country of origin/ancestry, n (%)						
India	567 (74.7%)	160 (63.2%)		175 (5.2%)	19 (4.3%)	
Pakistan	163 (21.5%)	74 (29.2%)	0.001	3198 (94.2%)	425 (95.5%)	0.37 ^c
Other	29 (3.8%)	19 (7.5%)		23 (0.7%)	1 (0.2%)	
Born in South Asia, n (%)	671 (88.5%)	225 (88.9%)	0.95	1836 (54.2%)	291 (65.7%)	6.50×10^{-06}
Years in recruitment country (Canada/UK) ^d	6.4 (5.8)	7.4 (5.8)	0.02	9.3 (9)	12.1 (9.4)	3.88×10^{-06}
Parental history of diabetes, n (%)	282 (37.3%)	142 (56.1%)	2.25×10^{-07}	891 (27.4%)	170 (38.9%)	8.88×10^{-07}
Vegetarians, n (%)	266 (37%)	84 (34.6%)	0.54	12 (1.3%)	1 (1.1%)	> 0.99 ^c
Low diet quality, n (%)	180 (24)	88 (35.1)	8.00×10^{-04}	-	-	-
Polygenic risk score (z-scores)	-0.11 (1)	0.347 (0.93)	1.51×10^{-08}	-0.04 (0.99)	0.32 (1.04)	4.98×10^{-12}
Polygenic risk score						
Tertile 1	240 (37.7%)	39 (19.4%)		1309 (34.4%)	117 (24.3%)	
Tertile 2	206 (32.4%)	73 (36.3%)	2.74×10^{-06}	1291 (33.9%)	142 (29.5%)	7.60×10^{-10}
Tertile 3	190 (29.9%)	89 (44.3%)		1209 (31.7%)	222 (46.2%)	
Fasting plasma glucose, mmol/L	4.27 (0.32)	5.02 (0.83)	5.51×10^{-32}	4.53 (0.41)	5.34 (1.14)	3.18×10^{-43}

1h post-load glucose, mmol/L	7.31 (1.38)	10.26 (2.02)	6.04×10^{-57}	-	-	-
2h post-load glucose, mmol/L	5.96 (1.16)	8.47 (2.16)	1.53×10^{-42}	5.49 (1.02)	9.14 (1.97)	1.57×10^{-155}
Area under curve glucose, mmol.h ^e	12.43 (1.83)	17.02 (2.89)	2.27×10^{-63}	10.02 (1.21)	14.48 (2.77)	3.82×10^{-133}

Characteristics of participants with available PRS and GDM_{IADPSG}, FPG, 1h, 2h post-load glucose levels or AUC glucose data. Presented data are means (Standard Deviation) unless otherwise indicated. P-Values are calculated from Chi-squared test for categorical variables and independent t-test for continuous variables. ^a pre-pregnancy values in START vs. weight at antenatal clinic (average 12 completed weeks of pregnancy) in BiB. ^b Derived using height measured at initial visit (in both studies) and pre-pregnancy weights (START) or antenatal clinic weights (BiB). ^c approximation may be incorrect due to small counts. ^d Canada for START samples and UK for BiB. ^e derived using fasting, 1h and 2h post-load measurements in START vs. fasting and 2h post-load measurements in BiB. Abbreviations: AUC, area under the curve glucose; BiB, Born in Bradford; BMI, body mass index; GDM, gestational diabetes mellitus; IADPSG, International Association of Diabetes and Pregnancy Study Groups; START, south Asian birth cohort; T2D, type 2 diabetes; UK, United Kingdom; vs. versus.

Table 2: Association between GDM risk factors and GDM related traits: results from multivariate models in START and BiB cohorts.

Dependent Variable	Independent Variables	START		BiB		Meta-analysis			
		Beta/OR [95%CI]	P-value	Beta/OR [95%CI]	P-value	Beta (SE) / OR [95%CI]	P-value	I ²	Q _E P-value
Fasting glucose	PRS (per 1SD increase)	0.083 [0.043 - 0.123]	6.00×10^{-05}	0.085 [0.065 - 0.105]	1.67×10^{-16}	0.085 [0.067 - 0.103]	2.85×10^{-20}	0	0.92
	Age (year)	0.021 [0.01 - 0.032]	2.00×10^{-04}	0.014 [0.009 - 0.019]	2.49×10^{-08}	0.015 [0.011 - 0.02]	4.19×10^{-11}	22	0.26
	BMI (kg/m ²)	0.024 [0.014 - 0.033]	5.51×10^{-07}	0.032 [0.028 - 0.036]	6.53×10^{-53}	0.031 [0.027 - 0.034]	7.99×10^{-60}	63	0.1
	Born in South Asia (Yes/No)	0.037 [-0.088 - 0.162]	0.56	0.08 [0.039 - 0.122]	2.00×10^{-04}	0.076 [0.037 - 0.115]	2.00×10^{-04}	0	0.52
	Parental history of T2D (Yes/No)	0.04 [-0.043 - 0.123]	0.34	0.066 [0.02 - 0.111]	0.005	0.06 [0.02 - 0.1]	0.003	0	0.6
	Parity	-0.046 [-0.102 - 0.01]	0.11	-0.015 [-0.033 - 0.004]	0.13	-0.018 [-0.036 - 0]	0.05	9	0.29
	Education level (per level)	-0.031 [-0.068 - 0.006]	0.1	-0.016 [-0.035 - 0.002]	0.09	-0.019 [-0.036 - -0.002]	0.02	0	0.49
	Low diet quality (Yes/No)	0.102 [0.01 - 0.193]	0.03	-	-	-	-	-	-
2h postload glucose	PRS (per 1SD increase)	0.189 [0.068 - 0.311]	0.002	0.211 [0.156 - 0.266]	7.87×10^{-14}	0.207 [0.157 - 0.257]	5.49×10^{-16}	0	0.75
	Age (year)	0.127 [0.093 - 0.161]	8.34×10^{-13}	0.068 [0.055 - 0.082]	8.82×10^{-23}	0.076 [0.064 - 0.089]	1.49×10^{-32}	90	0.002
	BMI (kg/m ²)	0.047 [0.019 - 0.074]	0.001	0.064 [0.053 - 0.075]	1.43×10^{-29}	0.062 [0.051 - 0.072]	3.44×10^{-32}	23	0.25
	Born in South Asia (Yes/No)	0.298 [-0.08 - 0.675]	0.12	0.308 [0.195 - 0.422]	1.14×10^{-07}	0.308 [0.199 - 0.416]	3.09×10^{-08}	0	0.96
	Parental history of T2D (Yes/No)	0.361 [0.109 - 0.613]	0.005	0.242 [0.117 - 0.366]	1.00×10^{-04}	0.265 [0.154 - 0.377]	3.23×10^{-06}	0	0.4
	Parity	-0.279 [-0.45 - -0.109]	0.001	-0.095 [-0.146 - -0.043]	3.00×10^{-04}	-0.11 [-0.16 - -0.061]	1.00×10^{-05}	76	0.04
	Education level (per level)	-0.063 [-0.176 - 0.051]	0.28	-0.073 [-0.124 - -0.022]	0.005	-0.071 [-0.118 - -0.025]	0.002	0	0.87
	Low diet quality (Yes/No)	0.365 [0.086 - 0.644]	0.01	-	-	-	-	-	-
AUC glucose	PRS (per 1SD increase)	0.165 [0.099 - 0.231]	1.08×10^{-06}	0.152 [0.119 - 0.185]	2.50×10^{-19}	0.155 [0.125 - 0.184]	7.74×10^{-25}	0	0.74
	Age (per year)	0.068 [0.05 - 0.087]	1.16×10^{-12}	0.043 [0.035 - 0.051]	1.01×10^{-24}	0.047 [0.039 - 0.054]	3.34×10^{-35}	84	0.01
	BMI (kg/m ²)	0.047 [0.032 - 0.062]	2.12×10^{-09}	0.047 [0.041 - 0.054]	8.89×10^{-44}	0.047 [0.041 - 0.053]	4.27×10^{-53}	0	0.94
	Born in South Asia (Yes/No)	0.081 [-0.123 - 0.285]	0.44	0.201 [0.133 - 0.269]	8.15×10^{-09}	0.189 [0.124 - 0.253]	1.01×10^{-08}	17	0.27
	Parental history of T2D (Yes/No)	0.122 [-0.015 - 0.258]	0.08	0.138 [0.063 - 0.213]	3.00×10^{-04}	0.134 [0.069 - 0.2]	6.00×10^{-05}	0	0.83
	Parity	-0.122 [-0.214 - -0.029]	0.01	-0.057 [-0.088 - -0.026]	3.00×10^{-04}	-0.063 [-0.093 - -0.034]	2.00×10^{-05}	41	0.19
	Education level (per level)	-0.045 [-0.106 - 0.016]	0.15	-0.045 [-0.075 - -0.014]	0.004	-0.045 [-0.072 - -0.017]	0.001	0	0.99
	Low diet quality (Yes/No)	0.215 [0.064 - 0.366]	0.005	-	-	-	-	-	-
	PRS (per 1SD increase)	1.56 [1.3 - 1.88]	2.97×10^{-06}	1.42 [1.27 - 1.59]	1.09×10^{-09}	1.45 [1.32 - 1.6]	2.27×10^{-14}	0	0.4
	Age (year)	1.13 [1.07 - 1.19]	2.50×10^{-06}	1.1 [1.07 - 1.13]	1.07×10^{-13}	1.11 [1.08 - 1.13]	1.98×10^{-18}	0	0.4

GDM (IADPSG criteria)	BMI (kg/m ²)	1.08 [1.04 - 1.12]	1.00×10^{-04}	1.08 [1.06 - 1.11]	1.01×10^{-14}	1.08 [1.06 - 1.1]	6.25×10^{-18}	0	0.87
	Born in South Asia (Yes/No)	1.35 [0.78 - 2.43]	0.3	1.72 [1.35 - 2.19]	1.00×10^{-05}	1.65 [1.33 - 2.06]	8.37×10^{-06}	0	0.44
	Parental history of T2D (Yes/No)	1.67 [1.17 - 2.38]	0.005	1.53 [1.21 - 1.94]	5.00×10^{-04}	1.57 [1.29 - 1.92]	7.06×10^{-06}	0	0.69
	Parity	0.86 [0.68 - 1.09]	0.23	0.87 [0.79 - 0.95]	0.003	0.87 [0.79 - 0.95]	0.001	0	0.99
	Education level (per level)	0.89 [0.76 - 1.05]	0.18	0.9 [0.81 - 0.99]	0.04	0.9 [0.82 - 0.98]	0.01	0	0.96
	Low diet quality (Yes/No)	1.68 [1.14 - 2.47]	0.008	-	-	-	-	-	-

Models were additionally adjusted for the first 5 PCs of each study. Abbreviations: BiB, Born in Bradford; BMI, Body mass index; CI, Confidence interval; GDM, Gestational diabetes mellitus; IADPSG, International Association of Diabetes and Pregnancy Study Groups; OR, Odds ratio; SA, South Asia; SD, Standard deviation; Q_E P, P-value from the test for (residual) heterogeneity; START, South Asian birth cohort; T2D, Type 2 diabetes.

Table 3: Population attributable fractions of GDM risk factors in mothers from the START and Born in Bradford studies (multivariable models).

Independent Variable	START		BiB		Meta-analysis			
	AF [95%CI]	P-value	AF [95%CI]	P-value	AF [95%CI]	P-value	I ²	Q _E P-value
Age (29-31 vs. <29 years)	5.6 [-9.1 - 20.2]	0.46	8.3 [3.5 - 13]	6.00×10^{-04}	8 [3.5 - 12.5]	5.00×10^{-04}	0	0.73
Age (>32yr vs. <29 years)	31.2 [17.1 - 45.3]	1.00×10^{-05}	20.2 [14.8 - 25.7]	4.72×10^{-13}	21.7 [16.6 - 26.8]	9.19×10^{-17}	50	0.16
Body mass index (≥ 23 vs. < 23)	21.8 [8.7 - 34.9]	0.001	33.8 [25.4 - 42.2]	2.47×10^{-15}	30.3 [23.3 - 37.4]	3.59×10^{-17}	56	0.13
Born in SA (Yes vs. No)	13.5 [-17.2 - 44.3]	0.39	19.3 [12.6 - 26]	1.47×10^{-08}	19 [12.5 - 25.6]	1.07×10^{-08}	0	0.72
Education (Post-secondary vs. less)	-18.2 [-46.8 - 10.5]	0.21	-0.8 [-4.6 - 3.1]	0.7	-1.1 [-4.9 - 2.7]	0.58	28	0.24
Parental history of T2D (Yes vs. No)	15.1 [4.4 - 25.7]	0.005	8.3 [4.1 - 12.5]	1.00×10^{-04}	9.2 [5.3 - 13.1]	3.54×10^{-06}	26	0.24
PRS (Tertile 3 vs. 1+2)	13.8 [4.9 - 22.6]	0.002	12.2 [7.8 - 16.6]	5.14×10^{-08}	12.5 [8.6 - 16.5]	4.47×10^{-10}	0	0.76
Low Diet Quality (Yes vs. No)	8.9 [1.5 - 16.4]	0.02	-	-	-	-	-	-
Sum PAF of PRS (T3) and parental history of diabetes	28.9		20.5		21.7			

GDM status derived using IADPSG criteria. Multivariate models included age, BMI, region of birth (South Asia vs other), education, parental history of diabetes, parity, principal components 1 to 5, and diet quality (START only) when applicable. Abbreviations: BiB, Born in Bradford; BMI, Body mass index, CI, Confidence interval; GDM, Gestational diabetes mellitus; IADPSG, International Association of Diabetes and Pregnancy Study Groups; PAF, Population attributable fraction; PRS, Polygenic risk score; Q_E P, P-value from the test for (residual) heterogeneity; START, South Asian birth cohort.

Table 4: Interaction effects between GDM risk factors and T2D PRS in START and BiB.

Dependent Variable	Interaction term	START		BiB		Meta-analysis			
		Beta/OR [95% CI] ^a	P _{interaction}	Beta/OR [95% CI] ^a	P _{interaction}	Beta/OR [95% CI] ^a	P _{interaction}	I ²	Q _E P-value
Fasting glucose	PRS x Age	-0.006 [-0.016 - 0.003]	0.2	0.004 [0 - 0.008]	0.07	0.002 [-0.001 - 0.006]	0.23	72	0.06
	PRS x BMI	-0.01 [-0.019 - -0.002]	0.01	0.004 [0 - 0.008]	0.05	0.001 [-0.002 - 0.005]	0.42	89	0.002
	PRS x Born in South Asia	-0.137 [-0.268 - -0.006]	0.04	0.037 [-0.003 - 0.078]	0.07	0.022 [-0.016 - 0.061]	0.26	84	0.01
	PRS x Parental history of T2D	-0.059 [-0.139 - 0.022]	0.15	0.016 [-0.028 - 0.061]	0.48	-0.001 [-0.04 - 0.037]	0.94	61	0.11
	PRS x Parity	-0.014 [-0.062 - 0.034]	0.56	0.004 [-0.01 - 0.018]	0.6	0.002 [-0.011 - 0.016]	0.73	0	0.48
	PRS x Education level	-0.014 [-0.05 - 0.022]	0.45	-0.005 [-0.022 - 0.013]	0.6	-0.006 [-0.022 - 0.009]	0.42	0	0.65
	PRS x Low diet quality	0.141 [0.053 - 0.228]	0.002	-	-	-	-	-	-
2h post-load glucose	PRS x Age	0 [-0.03 - 0.03]	0.98	0.01 [-0.001 - 0.021]	0.07	0.009 [-0.001 - 0.019]	0.08	0	0.54
	PRS x BMI	-0.022 [-0.047 - 0.003]	0.09	0 [-0.01 - 0.011]	0.94	-0.003 [-0.012 - 0.007]	0.56	61	0.11
	PRS x Born in South Asia	-0.191 [-0.586 - 0.205]	0.34	0.072 [-0.039 - 0.182]	0.2	0.053 [-0.054 - 0.159]	0.33	36	0.21
	PRS x Parental history of T2D	-0.092 [-0.335 - 0.151]	0.46	0.055 [-0.066 - 0.177]	0.37	0.026 [-0.083 - 0.135]	0.64	11	0.29
	PRS x Parity	-0.039 [-0.184 - 0.107]	0.6	0.009 [-0.03 - 0.047]	0.66	0.005 [-0.032 - 0.043]	0.77	0	0.54
	PRS x Education level	0.037 [-0.072 - 0.146]	0.51	0.008 [-0.039 - 0.056]	0.73	0.013 [-0.031 - 0.056]	0.56	0	0.64
	PRS x Low diet quality	0.068 [-0.199 - 0.335]	0.62	-	-	-	-	-	-
AUC glucose	PRS x Age	-0.007 [-0.023 - 0.009]	0.41	0.004 [-0.002 - 0.011]	0.19	0.003 [-0.003 - 0.009]	0.36	37	0.21
	PRS x BMI	-0.014 [-0.027 - 0]	0.05	0.002 [-0.004 - 0.008]	0.52	-0.001 [-0.006 - 0.005]	0.82	77	0.04
	PRS x Born in South Asia	-0.126 [-0.34 - 0.088]	0.25	0.015 [-0.051 - 0.081]	0.65	0.003 [-0.06 - 0.066]	0.93	35	0.22
	PRS x Parental history of T2D	-0.027 [-0.158 - 0.105]	0.69	0.025 [-0.048 - 0.098]	0.49	0.013 [-0.051 - 0.077]	0.68	0	0.5
	PRS x Parity	-0.057 [-0.135 - 0.022]	0.16	0.006 [-0.017 - 0.029]	0.6	0.001 [-0.021 - 0.023]	0.91	56	0.13
	PRS x Education level	0.007 [-0.052 - 0.066]	0.82	-0.008 [-0.036 - 0.021]	0.6	-0.005 [-0.03 - 0.021]	0.71	0	0.67
	PRS x Low diet quality	0.07 [-0.074 - 0.214]	0.34	-	-	-	-	-	-
	PRS x Age	0.99 [0.94-1.03]	0.59	0.99 [0.96-1.01]	0.17	0.99 [0.97-1]	0.14	0	0.92

GDM (IADPSG criteria)	PRS x BMI	0.97 [0.94-1.01]	0.15	0.98 [0.96-1]	0.03	0.98 [0.96-0.99]	0.01	0	0.76
	PRS x Born in South Asia	0.65 [0.33-1.23]	0.2	1.04 [0.82-1.31]	0.76	0.98 [0.79-1.23]	0.89	41	0.19
	PRS x Parental history of T2D	0.72 [0.5-1.04]	0.08	1.07 [0.85-1.35]	0.59	0.95 [0.78-1.16]	0.63	67	0.08
	PRS x Parity	0.88 [0.71-1.09]	0.23	0.99 [0.92-1.07]	0.86	0.98 [0.91-1.05]	0.57	15	0.28
	PRS x Education level	1.01 [0.85-1.19]	0.93	0.93 [0.85-1.02]	0.12	0.95 [0.87-1.03]	0.2	0	0.4
	PRS x Low diet quality	1.26 [0.85-1.89]	0.26	-	-	-	-	-	-

Results from models adjusted for age, BMI, education level, birth region (South Asia vs. other), parity, parental history of diabetes, and genetic PC axes 1 to 5. ^aValues are Beta for continuous dependent variables (fasting 2h, and AUC glucose), and OR for binary variable (i.e. GDM). Abbreviations: AUC, area under the curve; BiB, Born in Bradford; BMI, body mass index; CI, confidence interval; GDM, gestational diabetes mellitus; IADPSG, International Association of Diabetes and Pregnancy Study Groups; OR, odds ratio; PRS, polygenic risk score; Q_E P, P-value from the test for (residual) heterogeneity; START, South Asian birth cohort.