

Low birthweight and risk of type 2 diabetes: a Mendelian randomisation study

Tiange Wang^{1,2,3}  · Tao Huang^{1,2} · Yanping Li² · Yan Zheng² · JoAnn E. Manson^{4,5,6} · Frank B. Hu^{2,6} · Lu Qi^{1,2,6}

Received: 27 March 2016 / Accepted: 27 May 2016 / Published online: 23 June 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract

Aims/hypothesis Low birthweight has been associated with a high risk of type 2 diabetes mellitus in observational studies. However, it remains unclear whether this relation is causal.

Methods The present study included 3627 individuals with type 2 diabetes and 12,974 control participants of European ancestry from the Nurses' Health Study and the Health Professionals Follow-Up Study. A genetic risk score (GRS) was calculated based on five low-birthweight-related single nucleotide polymorphisms (SNPs). We assessed the evidence

for causality first by examining the association of the GRS and the individual SNPs with type 2 diabetes, and second by performing a Mendelian randomisation analysis to estimate the potentially causal effect size of low birthweight on type 2 diabetes.

Results In a meta-analysis of the two studies, each 1 point increment in the GRS was associated with a 6% (95% CI 3%, 9%) higher risk of type 2 diabetes. *CCNLI* rs900400 and 5q11.2 rs4432842 showed dose–response associations with risk of type 2 diabetes; the corresponding ORs and 95% CIs were 1.09 (1.03, 1.16) and 1.09 (1.02, 1.16), respectively. Furthermore, we observed an overall Mendelian randomisation OR of 2.94 (95% CI 1.70, 5.16; $p < 0.001$) for type 2 diabetes per 1 SD lower genetically determined birthweight.

Conclusions/interpretation A genetically lowered birthweight was associated with increased susceptibility to type 2 diabetes. Our findings support a potential causal relation between birthweight and risk of type 2 diabetes, providing new evidence to support the role of intrauterine exposures in the pathogenesis of type 2 diabetes.

Electronic supplementary material The online version of this article (doi:10.1007/s00125-016-4019-z) contains peer-reviewed but unedited supplementary material, which is available to authorised users.

✉ Lu Qi
lqi1@tulane.edu

- ¹ Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, 1440 Canal Street, New Orleans, LA 70112, USA
- ² Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA
- ³ Shanghai Institute of Endocrine and Metabolic Diseases, Rui Jin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, People's Republic of China
- ⁴ Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA
- ⁵ Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
- ⁶ Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

Keywords Birthweight · Mendelian randomisation · Type 2 diabetes

Abbreviations

GRS	Genetic risk score
GWAS	Genome-wide association studies
HPFS	Health Professionals Follow-up Study
MET	Metabolic equivalent
NHS	Nurses' Health Study
SNP	Single nucleotide polymorphism

Introduction

Type 2 diabetes has become a major global public health problem through an increasing burden of complicated morbidity and mortality [1–3]. It has been hypothesised that prenatal development might influence the susceptibility to type 2 diabetes in later life [4, 5]. Low birthweight, a widely used indicator of retarded fetal growth and intrauterine malnutrition, has been consistently related to an increased risk of type 2 diabetes in observational studies [6–9]. However, it remains unclear whether intrauterine growth plays a causal role in the development of type 2 diabetes, as it is difficult to fully eliminate confounding by socioeconomic status and lifestyle factors in observational studies.

Genetic association analysis is less likely to be affected by confounding and has been increasingly employed to inform causality [10]. A recent genome-wide association study (GWAS) in up to 69,308 individuals of European ancestry identified several genetic variants associated with birthweight [11]. Birthweight is a widely used surrogate for restricted intrauterine growth, which has been suggested to affect the risk of type 2 diabetes in later life. The genetic variants influencing birthweight may be intended to capture exposures that restrict intrauterine growth. We hypothesised that the exposures influencing intrauterine growth might lie in the causal pathway for susceptibility to type 2 diabetes. Although the intrauterine risks are unlikely to be measured, the exposures that affect intrauterine growth can be reflected in the birthweight. It is therefore possible to estimate the potential causal relation between intrauterine risks and type 2 diabetes by using the genetic variants related to birthweight.

In the present study, we first examined the evidence for causality by testing whether the genetic predisposition to low birthweight, which was evaluated by the GWAS-identified genetic variants, was associated with type 2 diabetes. We then performed a Mendelian randomisation analysis to estimate the possible causal effect size of low birthweight on type 2 diabetes.

Methods

Participants The Nurses' Health Study (NHS) is a prospective cohort study of 121,700 US female registered nurses aged 30–55 years at the study inception in 1976 [12]. Between 1989 and 1990, 32,826 women provided blood samples. The Health Professionals Follow-Up Study (HPFS) is a prospective cohort study of 51,529 US male health professionals aged 40–75 years at the study inception in 1986 [13]. Between 1993 and 1999, 18,159 men provided blood samples. In both cohorts, information about medical history and lifestyle has been collected biennially by self-administered questionnaires since inception. Both studies

were approved by the human research committee at the Brigham and Women's Hospital (Boston, MA, USA), and all participants provided written informed consent. For this analysis, we used 1986 as the baseline date for the NHS and HPFS. The present analysis included 5928 men and 10,673 women of European ancestry with genotype data available based on previous GWASs [14–19].

Ascertainment of type 2 diabetes Diabetes was defined as self-reported diabetes, with information collected by a questionnaire and confirmed by a validated supplementary questionnaire [20, 21]. For the diagnosis of diabetes before 1998, we used the National Diabetes Data Group criteria to define diabetes [22], which included one of the following: one or more classic symptoms (excessive thirst, polyuria, weight loss, hunger, pruritus or coma) plus a fasting plasma glucose level of ≥ 7.8 mmol/l, a random plasma glucose level of ≥ 11.1 mmol/l or an OGTT 2 h plasma glucose level of ≥ 11.1 mmol/l; at least two elevated plasma glucose levels on different occasions in the absence of symptoms; or treatment with hypoglycaemic medication (insulin or an oral hypoglycaemic agent). For diagnosis of diabetes from 1998 onwards, we used the ADA diagnostic criteria [23]. These criteria were the same as those proposed by the National Diabetes Data Group, except for the elevated fasting plasma glucose criterion, for which the cut-off point was changed from 7.8 mmol/l to 7.0 mmol/l.

Participants with diagnosed type 2 diabetes from the cohort baseline to follow-up until 2012 for the NHS and 2010 for the HPFS were included as cases. The validity of self-reported diabetes was verified in two subsamples from the NHS and HPFS, respectively. A physician blinded to the information reported on the supplementary questionnaire reviewed the medical records according to the diagnostic criteria. The medical record review's confirmation rate of diabetes as reported by the supplementary questionnaire was 98% for the NHS and 97% for the HPFS [20, 21]. Control participants were defined as those free of diabetes through the follow-up.

Assessment of covariates Participants in the NHS and HPFS were requested to provide their birthweight on the 1992 and 1994 questionnaires, respectively, within categories (in kg) of <2.26, 2.27–2.49, 2.50–3.15, 3.16–3.82, 3.83–4.4, ≥ 4.5 and unknown in the NHS, and <2.50, 2.50–3.15, 3.16–3.82, 3.83–4.4, ≥ 4.5 and unknown in the HPFS. The validity of the self-reported birthweight data has previously been described [24, 25]. Self-reported birthweight was highly correlated with recorded birthweight ($r=0.74$) and the birthweight reported by the participants' mothers ($r=0.71$). BMI was calculated as the weight in kilograms divided by the square of height in metres. Information about smoking status and alcohol intake was derived from the baseline questionnaires [12, 13]. Physical activity was expressed as

metabolic equivalents (METs) per week using the reported time spent carrying out various activities, weighting each activity by its intensity level. The validity of the self-reported body weight and physical activity data has previously been described [26, 27]. Self-reported and measured weights were highly correlated at 0.97 for men and 0.97 for women [26].

Genotyping and genetic risk score calculation Single nucleotide polymorphism (SNP) genotyping and imputation have previously been described in detail [15]. In brief, samples were genotyped and analysed using the Affymetrix Genome-Wide Human SNP Array 6.0 (Santa Clara, CA, USA) and the Birdseed calling algorithm (www.broadinstitute.org/mpg/birdsuite/birdseed.html). All samples used in the present study achieved a call rate of >98%. We used MACH (www.sph.umich.edu/csg/abecasis/mach) to impute SNPs on chromosomes 1–22, with National Center for Biotechnology Information build 36 of phase II HapMap Centre d'Etude du Polymorphisme (CEU: Utah residents with northern and western European ancestry) data (release 22; <http://hapmap.ncbi.nlm.nih.gov/>) as the reference panel.

Seven SNPs have been identified as being associated with birthweight by a previous GWAS [11]. We excluded two SNPs lying in the *ADCY5* and *CDKALI* loci that have been implicated by GWASs in susceptibility to type 2 diabetes [28, 29], and calculated a genetic risk score (GRS) on the basis of the other five SNPs to estimate the genetic variation of low birthweight. The GRS was calculated using a weighted method according to each SNP's relative effect size (β coefficient), obtained from the GWAS data [11]. The calculation equation was: weighted $GRS = (\beta_1 \times SNP_1 + \beta_2 \times SNP_2 + \dots + \beta_5 \times SNP_5) \times (5 / \text{sum of the } \beta \text{ coefficients})$, where SNP_i is the number of risk alleles associated with low birthweight, coding as 0, 1 and 2.

Statistical analysis Statistical analyses were performed in SAS version 9.3 software (SAS Institute, Cary, NC, USA).

We assigned a median value to each category of birthweight to model this variable as a continuous variable in analyses. General linear models were applied to examine the relationships of the GRS and individual SNPs to birthweight and covariates. Quartiles of the GRS were categorised according to quartile cut-off points for the control participants. Comparisons of the proportions and mean values across quartiles of the GRS were calculated by χ^2 test and ANOVA, respectively. We tested the associations of the GRS and individual SNPs with type 2 diabetes by logistic regression models. Results from the NHS and HPFS were pooled using an inverse-variance-weighted, fixed-effect meta-analysis (all p for heterogeneity >0.05). A restricted cubic spline regression model, which can help to prevent problems resulting from inappropriate assumptions on

linearity, was used to test the linear relationship between the GRS (as the continuous variable) and the risk of type 2 diabetes [30].

We further conducted Mendelian randomisation analysis using the method previously described [10, 31, 32]. A schematic presentation of the Mendelian randomisation analysis is shown in Fig. 1. The β_3 estimates for each SNP to evaluate the association between low birthweight and risk of type 2 diabetes can be calculated from the direct measurements β_1 (the estimate of effect size of each SNP on birthweight) and β_2 (the estimate of effect size of each SNP on type 2 diabetes) as: $\beta_3 = \beta_2 / \beta_1$. The SE of β_3 is given by: $S_3 = \sqrt{\frac{1}{\beta_1^2 S_2^2}}$ where S_2 is the SE of β_2 . In this study, β_1 is the estimate of effect size of each SNP on birthweight standardised using z score transformation derived from the birthweight GWAS [11], and β_2 is the \log_e OR estimate of type 2 diabetes for each SNP calculated from the NHS and HPFS using an inverse-variance-weighted, fixed-effects meta-analysis (all p for heterogeneity >0.05). We first calculated β_3 estimates for each SNP at *CCNLI*, *LCORL*, *ADRB1* and *HMG2* loci, and on chromosome 5q11.2, and then combined β_3 estimates for all five SNPs analysed using inverse-variance-weighted, fixed-effects meta-analysis to obtain an overall estimate of genetically determined low birthweight associated with type 2 diabetes (all p for heterogeneity >0.05). In a Mendelian randomisation design, the overall β_3 estimate (\log_e OR) would be considered to be an estimate for the causal association between low birthweight and type 2 diabetes. The OR for type 2 diabetes associated with each 1 SD lower genetically determined birthweight can be given by $\exp(\text{overall } \beta_3)$. We also performed a two-sample Mendelian randomisation analysis using the latest genetic association summary statistics for the trans-ethnic type 2 diabetes GWAS meta-analysis [33].

Results

Characteristics of the participants at baseline The range of the GRS was 0–10 for both women and men; the mean value (SD) of the GRS was 4.45 (1.57) and 4.55 (1.57) among women and men, respectively. As expected, GRS was associated with birthweight in both women ($\beta = -0.014$ kg, $p = 0.001$) and men ($\beta = -0.018$ kg, $p = 0.001$; see electronic supplementary material [ESM] Fig. 1). The baseline characteristics of 16,601 participants from the NHS and HPFS according to quartiles of the low birthweight GRS are presented in Table 1. Mean values for age, BMI, alcohol intake, physical activity, total energy intake and proportions of current smokers did not differ across the quartiles of the GRS for either men or women (all $p > 0.05$). In addition, the GRS was not associated with waist circumference,

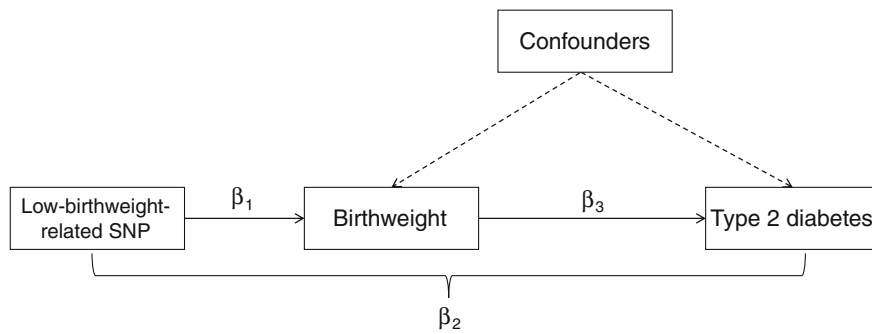


Fig. 1 Schematic representation of the Mendelian randomisation analysis. β_1 is the effect size estimate of each low-birthweight-related SNP on birthweight derived from the GWAS of birthweight reported by Horikoshi et al [11]. β_2 is the \log_e OR estimate of the type 2 diabetes for each SNP derived from the NHS and HPFS using inverse-variance-weighted, fixed-effects meta-analysis (all p for heterogeneity >0.05). β_3 is calculated from β_1 and β_2 for each SNP: $\beta_3 = \beta_2/\beta_1$. The SE of β_3 is given by: $S_3 = \sqrt{\frac{1}{\beta_1^2 S_2^2}}$ where S_2 is the SE of β_2 . We then combined β_3

estimates for all five SNPs analysed using inverse-variance-weighted, fixed-effects meta-analysis to obtain an overall estimate of the relation between genetically determined low birthweight and type 2 diabetes. In a Mendelian randomisation study, the overall β_3 estimate would be considered to be an estimate of the causal association between low birthweight and type 2 diabetes. The OR for type 2 diabetes associated with each 1 SD lower genetically determined birthweight can be given by $\exp(\text{overall } \beta_3)$

prevalent hypertension or hypercholesterolaemia at baseline (all $p > 0.05$; ESM Table 1).

Low-birthweight GRS and type 2 diabetes The present study included 3627 participants with and 12,974 participants without type 2 diabetes. The OR (95% CI) for type 2 diabetes per 1 SD lower birthweight was 1.35 (1.21, 1.50) and 1.24 (1.06, 1.45) in the NHS and HPFS, respectively.

As shown in Table 2, the association between each 1 point increment in the GRS and risk of type 2 diabetes was stronger in women (OR 1.09; 95% CI 1.05, 1.13) than men (OR 1.02; 95% CI 0.97, 1.07), with multivariable adjustment for age, BMI, smoking status, alcohol intake, physical activity, total energy intake and source of genotyping data. In a meta-analysis of the results from women and men, each 1 point increase in the GRS was

Table 1 Baseline characteristics among 5928 men and 10,673 women according to quartiles of the low-birthweight GRS

Variable	Quartiles of the GRS				<i>p</i> value
	Quartile 1	Quartile 2	Quartile 3	Quartile 4	
NHS					
<i>n</i>	2675	2273	3001	2724	–
Age (years)	54±6	54±6	54±6	54±6	0.61
BMI (kg/m ²)	25.8±4.9	25.7±4.6	25.8±4.7	25.8±4.8	0.67
Current smokers (%)	17.6	14.5	15.2	16.1	0.18
Alcohol (g/day)	6.4±10.2	6.6±10.7	6.4±10.1	6.2±9.8	0.61
Physical activity (MET-hr/wk)	13.8±15.1	13.8±16.4	13.7±15.9	14.4±17.2	0.46
Total energy intake (MJ/day)	7.46±2.03	7.50±2.01	7.47±2.05	7.44±2.02	0.74
GRS	2.5±0.7	3.8±0.3	4.9±0.4	6.5±0.8	<0.001
HPFS					
<i>n</i>	1492	1412	1529	1495	–
Age (years)	56±9	56±9	56±9	56±9	0.94
BMI (kg/m ²)	25.7±3.4	25.8±3.2	25.7±3.2	25.9±3.4	0.22
Current smokers (%)	8.8	9.2	8.7	8.6	0.76
Alcohol (g/day)	12.7±15.9	12.9±16.5	12.5±16.6	12.3±15.8	0.83
Physical activity (MET-hr/wk)	18.5±25.6	20.2±24.6	19.7±24.8	20.0±23.6	0.22
Total energy intake (MJ/day)	8.62±2.55	8.51±2.62	8.55±2.56	8.37±2.50	0.056
GRS	2.6±0.7	4.0±0.3	5.0±0.3	6.6±0.7	<0.001

Data are mean ± SD or percentage (%). *p* values for difference are based on ANOVA for continuous data or χ^2 test for categorical data

Table 2 Association between the low-birthweight GRS and type 2 diabetes

Variable	Continuous (per 1 point)	Quartiles of the GRSs				<i>p</i> for trend
		Quartile 1	Quartile 2	Quartile 3	Quartile 4	
NHS						
Cases (%)		19.4	21.0	21.8	22.3	
Model 1	1.05 (1.02, 1.09)	1.00	1.10 (0.96, 1.27)	1.16 (1.02, 1.32)	1.19 (1.04, 1.36)	0.007
Model 2	1.09 (1.05, 1.13)	1.00	1.19 (1.01, 1.40)	1.24 (1.06, 1.44)	1.32 (1.13, 1.54)	<0.001
HPFS						
Cases (%)		21.9	22.6	23.5	24.4	
Model 1	1.03 (0.99, 1.07)	1.00	1.04 (0.87, 1.24)	1.10 (0.93, 1.30)	1.15 (0.97, 1.36)	0.099
Model 2	1.02 (0.97, 1.07)	1.00	1.12 (0.91, 1.38)	1.10 (0.89, 1.35)	1.13 (0.92, 1.39)	0.242
Pooled data						
Model 1	1.04 (1.02, 1.07)	1.00	1.08 (0.96, 1.20)	1.14 (1.02, 1.25)	1.17 (1.05, 1.30)	0.002
Model 2	1.06 (1.03, 1.09)	1.00	1.16 (1.01, 1.31)	1.18 (1.03, 1.33)	1.24 (1.08, 1.39)	<0.001

Data are OR (95% CI) for type 2 diabetes; pooled data based on the results of the two studies using inverse-variance-weighted, fixed-effects meta-analysis (all *p* for heterogeneity >0.05)

Model 1: adjusted for age; model 2: adjusted for age, BMI, smoking, alcohol intake, physical activity, total energy intake and source of genotyping data

associated with a 6% (95% CI 3%, 9%) increased risk of type 2 diabetes; the OR (95% CI) for type 2 diabetes was 1.16 (1.01, 1.31), 1.18 (1.03, 1.33) and 1.24 (1.08, 1.39) for the second, third and fourth quartiles of the GRS, respectively, compared with the lowest quartile (*p* for trend <0.001). Combining the data for both women and men, the low-birthweight GRS showed a linear relation with an increased risk of type 2 diabetes (*p* for linearity=0.001; Fig. 2).

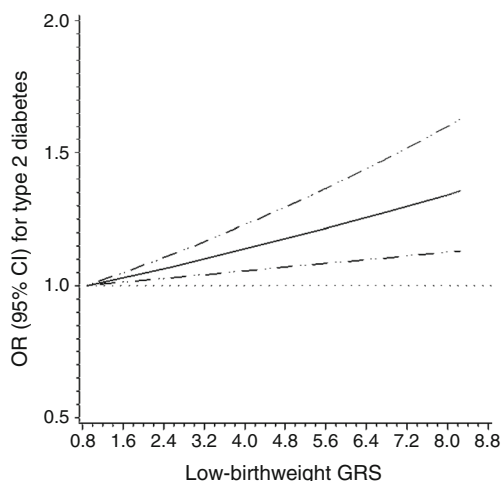


Fig. 2 Relationship between the low-birthweight GRS and type 2 diabetes. Data are ORs (solid lines) and 95% CIs (dashed lines), based on the combined data of the NHS and HPFS, adjusting for age and sex. *p* for linearity=0.001

Individual SNPs in relation to birthweight and type 2 diabetes Characteristics of individual SNPs and their relation to birthweight and type 2 diabetes are depicted in Table 3. The results of the NHS and HPFS were pooled using inverse-variance-weighted, fixed-effects meta-analysis (all *p* for heterogeneity >0.05). Of the five SNPs, *CCNLI* rs900400 had the strongest association with a lower birthweight (*p*<0.001); *CCNLI* rs900400 and 5q11.2 rs4432842 showed dose-response associations with risk of type 2 diabetes, with OR (95% CI) values of 1.09 (1.03, 1.16) and 1.09 (1.02, 1.16), respectively.

Mendelian randomisation analysis We further performed Mendelian randomisation analysis to estimate the causal effect size of genetically determined low birthweight on type 2 diabetes (Fig. 3). β_3 is the \log_e OR estimate of the association between genetically determined low birthweight and type 2 diabetes for each SNP. Of the five SNPs, *CCNLI* rs900400 and 5q11.2 rs4432842 showed statistically significant associations with type 2 diabetes (β_3 =1.21 and 2.56, respectively; both *p*<0.05). We then combined β_3 estimates for all five SNPs using inverse-variance-weighted, fixed-effects meta-analysis to obtain an overall β_3 estimate (*p* for heterogeneity=0.318). The overall β_3 estimate was 1.08 (95% CI 0.53, 1.64). By exponentiating the overall β_3 , we calculated an OR of 2.94 (95% CI 1.70, 5.16; *p*<0.001), indicating that each 1 SD lower genetically determined birthweight was associated with a 1.94 (95% CI 0.70, 4.16)-fold increased risk of type 2 diabetes.

We also derived data and performed a Mendelian randomisation analysis using the latest genetic association summary statistics for recent type 2 diabetes GWAS [33]. A

Table 3 Characteristics of the genetic variants and the association of increasing number of low-birthweight-related alleles with birthweight and type 2 diabetes

SNP	Gene	Chromosome	Effect allele/other ^a	Frequency of effect allele		$\beta \pm \text{SE}$ for birthweight ^b (kg)	OR (95% CI) for type 2 diabetes ^c
				NHS	HPFS		
rs900400	<i>CCNL1</i>	3	C/T	17.6	17.9	-0.028 ± 0.008	1.09 (1.03, 1.16)
rs724577	<i>LCORL</i>	4	C/A	54.8	55.9	-0.012 ± 0.009	1.04 (0.97, 1.11)
rs4432842	5q11.2	5	C/T	8.9	8.1	-0.002 ± 0.009	1.09 (1.02, 1.16)
rs1801253	<i>ADRB1</i>	10	G/C	7.5	9.1	-0.012 ± 0.009	1.00 (0.93, 1.06)
rs1042725	<i>HMGA2</i>	12	T/C	24.2	27.7	-0.014 ± 0.008	1.05 (0.99, 1.11)

Pooled data based on the results of the two studies using inverse-variance-weighted, fixed-effects meta-analysis (all p for heterogeneity >0.05)

^a Allele coding was based on the forward strand. The effect allele is associated with low birthweight, and the other allele is the reference allele

^b The β coefficient represents the coefficients of birthweight (kg) associated with each increment in the low-birthweight-related allele for genetic variants, adjusted for source of genotyping data

^c OR represents the risk of type 2 diabetes associated with each increment in the low-birthweight-related allele for genetic variants, adjusted for age and source of genotyping data

meta-analysis of the associations between the five SNPs and type 2 diabetes showed an OR of 1.03 (CI 1.01, 1.04); each 1 SD lower genetically determined birthweight showed a Mendelian randomisation OR of 1.70 (95% CI 1.32, 2.19; $p < 0.001$) for type 2 diabetes (ESM Table 2).

Discussion

We tested the potential causal effect of low birthweight on type 2 diabetes in two large prospective cohorts of US men and women. We found an association between the overall genetic susceptibility to low birthweight, which was estimated by a GRS, and an increased risk of type 2 diabetes. We further applied the Mendelian randomisation approach to provide

supportive evidence for the causality, and the findings were confirmed using the summary statistics from the latest GWAS.

Evidence from both population and experimental studies has suggested that restricted early life development has a long-term structural and functional influence on individuals' predisposition to an increased risk of metabolic diseases such as type 2 diabetes. Famine studies from the Netherlands, China and Ukraine have shown that individuals with retarded intrauterine growth are more susceptible to type 2 diabetes later in life [34–36]. As a widely used marker of fetal malnutrition and growth retardation, low birthweight has been consistently associated with a high risk of type 2 diabetes. A meta-analysis of 14 studies showed a U-shaped relation between birthweight and risk of type 2 diabetes [6]. A more recent larger meta-analysis of 31 studies confirmed the relation but suggested that the overall relation between

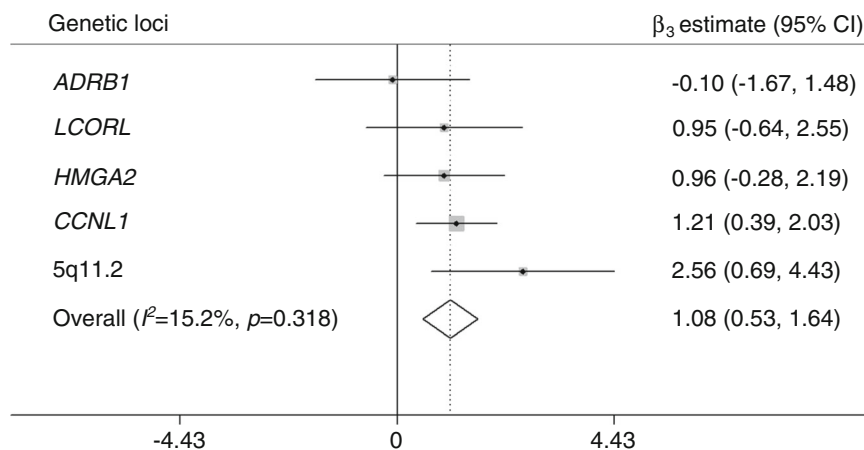


Fig. 3 Mendelian randomisation estimate of the association of low birthweight with risk of type 2 diabetes. The forest plot shows β_3 estimates (\log_e -ORs) of the effect of low birthweight on the risk of type 2 diabetes for each low-birthweight-related SNP. The overall β_3 estimate

was obtained by using inverse-variance-weighted, fixed-effects meta-analysis (p for heterogeneity = 0.318), and can be interpreted as an OR of 2.94 (95% CI 1.70, 5.16) for type 2 diabetes per 1 SD lower genetically determined birthweight ($p < 0.001$)

birthweight and type 2 diabetes was likely to be linear, i.e. that lower birthweight is related to an increased risk of type 2 diabetes [7]. In the present study, we also reported a linear relation between the low-birthweight GRS and type 2 diabetes, which was in line with the previous observational findings. Although the genetic association in men appeared to be weaker than that in women, the sex difference was not statistically significant ($p > 0.05$).

According to Mendel's second law, the alleles of one gene sort into gametes independently of the alleles of another gene. Therefore, genetic associations are less likely to be affected by confounding and have been increasingly used in causal inference [10, 31, 32]. Based on the positive findings of genetic associations, we further used a Mendelian randomisation approach, which can provide an unconfounded estimate of a causal relationship between exposure and disease outcome [37], to test the possible causality. Notably, birthweight itself may not be an exposure relevant to type 2 diabetes; instead, the exposures influencing intrauterine growth are more likely to play a causal role in the development of the disease. As an indicator of intrauterine growth, birthweight reflects the influence of intrauterine exposures on fetal growth. Using the genetic variants related to birthweight, our results lent support to a causal relation between low birthweight and increased risk of type 2 diabetes. Interestingly, the findings of the Mendelian randomisation analysis suggested that the association between genetically determined low birthweight and type 2 diabetes was mainly driven by the SNPs at the *CCNLI* locus and on chromosome 5q11.2, which have so far not been related to any known risk factors for diabetes. Therefore, the findings are less likely to be affected by the pleiotropic effects of the genetic variants. Moreover, we performed a two-sample Mendelian randomisation analysis using the summary statistics for the latest type 2 diabetes GWAS, and both the summarised genetic association and Mendelian analysis results supported a potential causal relation.

The major strengths of the present study include the prospective design, the high-quality genetic data, the utilisation of genetic and Mendelian randomisation approaches and replication in well-powered summary statistics GWAS data. To our knowledge, this study is the first to investigate the potential causal relation between low birthweight and risk of type 2 diabetes. We also acknowledge several limitations. First, the study included only white participants, and future investigations in other ethnic populations are needed to verify our findings. Second, although we excluded the likely pleiotropic variants at the *CDKAL1* and *ADCY5* loci from the analysis, little is known about the mechanisms underlying the other five loci. It is possible that some or all of these loci could also influence the processes leading to type 2 diabetes independently of intrauterine growth. Third, we used individual SNPs rather

than a GRS as instrumental variables in the Mendelian randomisation analysis, and the potential low power and weak instrument bias could affect the results. Moreover, as the Mendelian randomisation analysis was relatively small, future studies with larger sample sizes are warranted to confirm the findings.

In conclusion, we demonstrated that genetically lowered birthweight was associated with an increased risk of type 2 diabetes. Our findings validate the epidemiological observation of an inverse association between birthweight and type 2 diabetes, and provides new evidence of a role for intrauterine exposures in the pathogenesis of type 2 diabetes.

Acknowledgements We are indebted to the participants of the NHS and HPFS for their continuing outstanding support, and colleagues working in this study for their valuable help.

Funding The cohorts were supported by grants P01 CA49449, UM1 CA186107 and UM1 CA167552 from the National Institutes of Health. This work was supported by grants from the National Heart, Lung, and Blood Institute (HL071981, HL034594, HL126024), the National Institute of Diabetes and Digestive and Kidney Diseases (DK091718, DK100383, DK078616) and United States–Israel Binational Science Foundation Grant 2011036. LQ was a recipient of the American Heart Association Scientist Development Award (0730094 N).

Duality of interest The authors declare that there is no duality of interest associated with this manuscript.

Contribution statement LQ was involved in the conception and design, provision of study materials and patients and collection and assembly of data, and also obtained funding for the study. TW, TH, YL, YZ, JEM, FBH and LQ contributed to the analysis and interpretation of the data. TW produced a draft of the manuscript. TW, TH, YL, YZ, JEM, FBH and LQ were involved in the critical revision of the manuscript for important intellectual content. TW, TH, YL, YZ, JEM, FBH and LQ approved the final manuscript. LQ is responsible for the integrity of the work as a whole.

References

1. Danaei G, Finucane MM, Lu Y et al (2011) National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. *Lancet* 378:31–40
2. Zimmet P, Alberti KG, Shaw J (2001) Global and societal implications of the diabetes epidemic. *Nature* 414:782–787
3. Frayling TM (2007) Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nat Rev Genet* 8:657–662
4. Bateson P, Barker D, Clutton-Brock T et al (2004) Developmental plasticity and human health. *Nature* 430:419–421
5. Gluckman PD, Hanson MA, Cooper C, Thornburg KL (2008) Effect of in utero and early-life conditions on adult health and disease. *N Engl J Med* 359:61–73
6. Harder T, Rodekamp E, Schellong K, Dudenhausen JW, Plagemann A (2007) Birth weight and subsequent risk of type 2 diabetes: a meta-analysis. *Am J Epidemiol* 165:849–857

7. Whincup PH, Kaye SJ, Owen CG et al (2008) Birth weight and risk of type 2 diabetes: a systematic review. *JAMA* 300:2886–2897
8. Pilgaard K, Færch K, Carstensen B et al (2010) Low birthweight and premature birth are both associated with type 2 diabetes in a random sample of middle-aged Danes. *Diabetologia* 53:2526–2530
9. Ruiz-Narváez EA, Palmer JR, Gerlovin H et al (2014) Birth weight and risk of type 2 diabetes in the Black Women's Health Study: does adult BMI play a mediating role? *Diabetes Care* 37:2572–2578
10. Nelson CP, Hamby SE, Saleheen D et al (2015) Genetically determined height and coronary artery disease. *N Engl J Med* 372:1608–1618
11. Horikoshi M, Yaghoobkar H, Mook-Kanamori DO et al (2013) New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. *Nat Genet* 45:76–82
12. Colditz GA, Manson JE, Hankinson SE (1997) The Nurses' Health Study: 20-year contribution to the understanding of health among women. *J Womens Health* 6:49–62
13. Rimm EB, Giovannucci EL, Willett WC et al (1991) Prospective study of alcohol consumption and risk of coronary disease in men. *Lancet* 338:464–468
14. Hunter DJ, Kraft P, Jacobs KB et al (2007) A genome-wide association study identifies alleles in *FGFR2* associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 39:870–874
15. Qi L, Cornelis MC, Kraft P et al (2010) Genetic variants at 2q24 are associated with susceptibility to type 2 diabetes. *Hum Mol Genet* 19:2706–2715
16. Cornelis MC, Monda KL, Yu K et al (2011) Genome-wide meta-analysis identifies regions on 7p21 (*AHR*) and 15q24 (*CYP1A2*) as determinants of habitual caffeine consumption. *PLoS Genet* 7, e1002033
17. Jensen MK, Pers TH, Dworzynski P, Girman CJ, Brunak S, Rimm EB (2011) Protein interaction-based genome-wide analysis of incident coronary heart disease. *Circ Cardiovasc Genet* 4:549–556
18. Köttgen A, Albrecht E, Teumer A et al (2013) Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nat Genet* 45:145–154
19. Qi Q, Chu AY, Kang JH et al (2012) Sugar-sweetened beverages and genetic risk of obesity. *N Engl J Med* 367:1387–1396
20. Manson JE, Rimm EB, Stampfer MJ et al (1991) Physical activity and incidence of non-insulin-dependent diabetes mellitus in women. *Lancet* 338:774–778
21. Hu FB, Leitzmann MF, Stampfer MJ, Colditz GA, Willett WC, Rimm EB (2001) Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. *Arch Intern Med* 161:1542–1548
22. National Diabetes Data Group (1979) Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes* 28:1039–1057
23. American Diabetes Association (1997) Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care* 20:1183–1197
24. Troy LM, Michels KB, Hunter DJ et al (1996) Self-reported birthweight and history of having been breastfed among younger women: an assessment of validity. *Int J Epidemiol* 25:122–127
25. Curhan GC, Willett WC, Rimm EB, Spiegelman D, Ascherio AL, Stampfer MJ (1996) Birth weight and adult hypertension, diabetes mellitus, and obesity in US men. *Circulation* 94:3246–3250
26. Rimm EB, Stampfer MJ, Colditz GA, Chute CG, Litin LB, Willett WC (1990) Validity of self-reported waist and hip circumferences in men and women. *Epidemiology* 1:466–473
27. Wolf AM, Hunter DJ, Colditz GA et al (1994) Reproducibility and validity of a self-administered physical activity questionnaire. *Int J Epidemiol* 23:991–999
28. Steinthorsdottir V, Thorleifsson G, Reynisdottir I et al (2007) A variant in *CDKAL1* influences insulin response and risk of type 2 diabetes. *Nat Genet* 39:770–775
29. Dupuis J, Langenberg C, Prokopenko I et al (2010) New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* 42:105–116
30. Durrleman S, Simon R (1989) Flexible regression models with cubic splines. *Stat Med* 8:551–561
31. International Consortium for Blood Pressure Genome-Wide Association Studies, Ehret GB, Munroe PB et al (2011) Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* 478:103–109
32. Mokry LE, Ross S, Ahmad OS et al (2015) Vitamin D and risk of multiple sclerosis: a Mendelian randomization study. *PLoS Med* 12, e1001866
33. DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium; Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium et al (2014) Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet* 46:234–244
34. Hales CN, Barker DJ, Clark PM et al (1991) Fetal and infant growth and impaired glucose tolerance at age 64. *BMJ* 303:1019–1022
35. Li Y, He Y, Qi L et al (2010) Exposure to the Chinese famine in early life and the risk of hyperglycemia and type 2 diabetes in adulthood. *Diabetes* 59:2400–2406
36. Lumey LH, Khalangot MD, Vaiserman AM (2015) Association between type 2 diabetes and prenatal exposure to the Ukraine famine of 1932–33: a retrospective cohort study. *Lancet Diabetes Endocrinol* 3:787–794
37. Davey Smith G, Ebrahim S (2003) 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 32:1–22