2020-03-01

Genetic Susceptibility Determines -Cell Function and Fasting Glycemia Trajectories Throughout Childhood: A 12-Year Cohort Study (EarlyBird 76)

Carayol, J

10.2337/dc19-0806
Diabetes Care
American Diabetes Association

Genetic susceptibility determines β-cell function and fasting glycaemia trajectories throughout childhood: a 12-yr cohort study (EARLYBIRD 76).

Jerome Carayol[1*], Joanne Hosking[2], Jonathan Pinkney[2], Julien Marquis[1,3], Aline Charpagne[1,4], Sylviane Metairon[1], Alison Jeffery[2], Jörg Hager[1], Francois-Pierre Martin[1*]

[1] Nestlé Institute of Health Sciences, Nestlé Research, EPFL Innovation Park, CH-1015 Lausanne, Switzerland.

[2] Plymouth University, Faculty of Medicine and Dentistry, Plymouth Science Park, 1 Davey Road, Plymouth PL6 8BX, United Kingdom.

[3] Present address: University of Lausanne, Faculty of biology and medicine, Genopode building, 1015 Lausanne, Switzerland.

[4] Present address: Sophia Genetics SA, rue du Centre 172, 1025, Saint Sulpice, Switzerland.

Corresponding authors:

Dr. Jerome Carayol, PhD, Nestlé Institute of Health Sciences, Nestlé Research, EPFL Innovation Park, Building G, CH-1015 Lausanne, (Switzerland), Jerome.carayol@rd.nestle.com

Dr. François-Pierre Martin, PhD, Nestlé Institute of Health Sciences, Nestlé Research, EPFL Innovation Park, Building H, CH-1015 Lausanne, (Switzerland), francois-pierre.martin@rd.nestle.com

Abstract

Objective: Previous studies suggested that childhood prediabetes may develop prior to obesity and be associated with relative insulin deficiency. We proposed that the insulin deficient phenotype is genetically determined, and tested this hypothesis by longitudinal modelling of insulin and glucose traits with diabetes risk genotypes in the EarlyBird cohort.

Research Design and Methods: EarlyBird is a non-intervention prospective cohort study that recruited 307 healthy UK children at age 5, and followed them throughout childhood. We genotyped 121 single nucleotide polymorphisms (SNPs) previously associated with diabetes risk, identified in the adult population. Association of SNPs with fasting insulin and glucose, HOMA-IR and HOMA-B, available from ages 5 to 16 years were tested. Association analysis with hormones was performed on selected SNPs.

Results: Several candidate loci influenced the course of glycemic and insulin traits, including rs780094 (GCKR), rs4457053 (ZBED3), rs11257655 (CDC123), rs12779790 (CDC123, CAMK1D), rs1111875 (HHEX), rs7178572 (HMG20A), rs9787485 (NRG3), and rs1535500 (KCNK16). Some of these SNPs interacted with age, growth hormone-IGF-1 axis, adrenal and sex steroid activity.

Conclusions: The findings that genetic markers influence both elevated and average courses of glycemic traits and β-cell function in children during puberty independently of BMI is a significant step towards early identification of children at risk of diabetes. These findings build on our previous observations that pancreatic β-cell defects predate insulin resistance in the onset of prediabetes. Understanding the mechanisms of interactions between genetic factors, puberty and weight gain would allow the development of new and earlier disease management strategies in children.

Key words: β-cell function, Children, Fasting glucose, Genetic susceptibility, Insulin resistance, Insulin secretion.

Introduction

Diabetes is now one of the most common non-communicable diseases in the world. The World Health Organization (WHO) reports that diabetes affects over 422 million people, or 8.5% of the world's adult population (1). It has been projected that one in every three individuals born in the US in the year 2000 will develop diabetes during their lifetime (2). However, the increasing burden of diabetes will be felt, according to WHO, most acutely in low and middle-income countries, with catastrophic economic consequences (1). Therefore, there is a pressing need to improve the prediction and early prevention of diabetes.

Diabetes results from impaired insulin secretion, resistance to the action of insulin, or a combination of both factors. Failure of insulin secretion predominates in type 1 diabetes (T1D), whereas resistance to insulin (IR) and relative insulin deficiency characterize type 2 diabetes (T2D). As a result of the rising prevalence of obesity, T2D has become increasingly common in children and adolescents (3). Adolescence is a period of high vulnerability to the onset of diabetes in children, as a result of rapid endocrine changes, increasingly accompanied by obesity, and on a variable background of genetic risk (4). Yet, very little is known about the molecular pathways linking genetic risk factors, weight gain and puberty to the risk of developing diabetes in adolescents. Understanding these pathways is of fundamental importance to understanding disease progression in children and how prediabetes could be detected at an earlier stage to enable preventive measures to be taken to tackle the worldwide epidemic of diabetes (5).

The EarlyBird study is a landmark prospective cohort study investigating the origins of T2D in children. This cohort of healthy children has been followed from ages 5 to 16 years with annual clinical, anthropometric and physiological measurements (6). We were amongst the first to report the occurrence of an early defect in β-cell function among children who go on to develop

prediabetes (6). Some 17% of the nominally healthy children in the EarlyBird cohort showed impaired fasting glycemia (IFG) by the age of 15. Furthermore, children who developed IFG already exhibited higher fasting blood glucose levels at 5 years of age, compared with those who did not subsequently develop IFG, and this effect was independent of BMI (6). However, prediabetes did not appear until puberty when IR was at its highest. This is consistent with the US National Health and Nutrition Examination Survey (NHANES) 2005–2006, which also included children of normal-weight, and suggested that the prevalence of prediabetes in adolescence was strongly influenced by IR (7).

Information from longitudinal studies of healthy weight children is required to determine how prediabetes results from the interaction of genetic risk factors, weight gain, changing levels of IR and other endocrine parameters during puberty, and other non-genetic risk factors. (8). Therefore, the present analysis of the EarlyBird cohort was designed to examine how genetic variants, puberty, and weight gain interact to influence insulin action and blood glucose levels, at an early age of high vulnerability to diabetes.

We hypothesized that genes and SNPs known to be associated with increased risks of T2D, would be associated with trajectories of fasting glucose, insulin, and HOMA indices of insulin resistance (HOMA-IR) and β-cell function (HOMA-B) in children. We also explored if insulin action may be influenced by genetic variations in the growth hormone-IGF-1 axis, adrenal and sex steroid activity. The aim of this study was to identify specific genes and SNPs associated with these glycemic traits independently of body weight that could identify young people at high risk of diabetes. The long term goal of this work is to develop risk-modifying interventions before adolescence.

RESEARCH DESIGN AND METHODS

Study design and participants

The study was conducted in accordance with the principles of the Declaration of Helsinki II. Ethical approval was granted by the Plymouth Local Research Ethics Committee (1999), and parents gave written consent and children verbal assent. The EarlyBird Diabetes Study incorporates a 1995/1996 birth cohort recruited in 2000/2001 when the children were 5 years old (307 children, 170 boys) (9). Most of the children were white Caucasian and five children were of mixed race, reflecting the racial mix of the city of Plymouth. According to the paediatric thresholds for overweight and obesity proposed by the International Obesity Task Force (IOTF), 13% of the EarlyBird boys and 26% of girls were overweight at the baseline, which included 4% and 5% respectively who were obese. The thresholds approximate to the 91st and 98th centiles of the 1990 BMI reference curves for the UK, and are deemed to correspond to equivalent thresholds in adulthood. At baseline, the number of children with family history of T2D were: mother (N=0), father (N=2), maternal grandmother (N=14), maternal grandfather (N=19), paternal grandmother (N=18), paternal grandfather (N=18). The collection of data from the EarlyBird cohort is composed of clinical and anthropometric variables measured on an annual basis from the age of 5 to 16.

Anthropometrics

BMI was derived from direct measurement of height (Leicester Height Measure; Child Growth Foundation, London, U.K.) and weight (Tanita Solar 1632 electronic scales), performed in duplicate and averaged. BMI SD scores were calculated from the British 1990 standards (10).

Laboratory assessment

The children were fasted overnight for 10 h before venesection. HOMA2IR and HOMA2B were determined each year from fasting glucose (Cobas Integra 700 analyzer; Roche Diagnostics) and insulin (DPC IMMULITE) (cross-reactivity with proinsulin,1%) using the homeostasis model

assessment program, which has been validated in children (11). Peripheral whole blood and serum was collected annually and stored at -80°C for analysis. Dehydroepiandrosterone sulfate (DHEAS), androstenedione, and testosterone were measured in serum by LC-MS/MS using the Waters Acquity Ultrahigh performance liquid chromatography system and Quattro Premier tandem quadrupole tandem mass spectrometry (Waters Corporation, MA). Free testosterone (FT) was calculated using the formula of Vermeulen (12). Serum sex hormone binding proteins (SHBG) was assayed using the Roche Cobas method on the E170 Modular Analytics system and IGF-1 was measured by chemiluminescence immunoassay (Nichols Institute Diagnostics, San Juan Capistrano, CA 92675, USA) using standards referenced to WHO 1st International Reference Reagent 1988 (Insulin-Like Growth Factor-1 87/518).

Genotyping

A total of 1793 associations with T2D are reported in the GWAS catalogue (https://www.ebi.ac.uk/gwas/) (13). Results from small studies (n<1000 subjects) were excluded. Only SNPs with genome-wide significance for T2D itself, HOMA-IR and response to metformin (but not T2D associated co-morbidities) and reported in European populations were selected. Excluding overlapping markers between those phenotypes the GWAS catalogue yields 116, 30 and 1 association respectively (147 SNPs). We were able to design genotyping assays for 136 of these SNPs that were included in the analysis. Descriptive information on selected SNPS are reported in Supplementary Table 1. Genomic DNA was extracted from blood, using the QIAsymphony DSP DNA Midi Kit (96) on a QIAsymphony automation platform (Qiagen). The DNA concentration was measured with a fluorimetric method (Picogreen, Thermo Fisher). Genotyping was performed using the SNPtype assay (Fluidigm) which relies on allele-specific PCR reactions that use three primers and two universal probes to distinguish between two alleles. Genotyping was performed using microfluidic chips requiring two chips per batch of 96 samples. Each batch was actually made

of 86 samples of interest, three non-template controls with pre-amplification, three non-template controls without pre-amplification, and two control samples of known genotype run in duplicate (Hapmap NA12891 and NA12892, obtained from the NIGMS Human Genetic Cell Repository at the Coriell Institute for Medical Research). DNA was normalized to 10 ng/uL, randomized, and 25 ng DNA was pre-amplified for 14 cycles with a pool of all 136 assay primer pairs using the QIAGEN Multiplex PCR Kit, following Fluidigm's recommendations. The amplification mixes were subsequently prepared from diluted pre-amplified products (1/100 in low TE buffer) and loaded together with the SNPtype assay mixes in 96.96 Dynamic Arrays IFC (Fluidigm). The final amplification and data collection were performed on a Biomark HD (Fluidigm). The analysis was performed with the Fluidigm SNP Genotyping Analysis software (version 4.1.2) which uses a cluster analysis method to automatically call genotypes (confidence threshold set at 65, SNPtype normalization, K-Means clustering method). Each SNPtype assay was manually quality controlled on each individual chip by evaluating the background fluorescence levels from non-template controls and the actual signal accuracy from the two positive controls. All genotypes were validated after manual inspection of each cluster, resulting in genetic data available for 121 SNPs. Quality control (QC) consisted in the exclusion of candidate SNPs with more than 5% missing values and subjects with more 10% missing values. The sample consisted of 318 children with clinical and genetic information for 121 SNPs after this QC. Pairwise linkage disequilibrium of SNPs was also calculated and reported in Supplementary Table 2. Comparison of minor allele frequencies using Earlybird genetic data, as well as 1000 Genome British in England and Scotland (GBR) population data, was conducted and reported in Supplementary Table 3.

Statistical Analysis

For insulin, glucose, HOMA-IR and HOMA-B parameters collected over the 12-years period (age 5 to 16 years old), outlier detection filtering was based on inter-quartile range (IQR). Values below

the lower bound (Q1 - 4 * IQR) or above the upper bound (Q3 + 4 * IQR) were set as missing for each period of age, where Q1 and Q2 correspond to the value for the 1st and 3rd quartile, respectively. Assuming subsequent repeated measurements analysis, subjects with missing data for at least 4 visits were excluded, resulting in the selection of 224 children.

Fasting insulin, HOMA-IR, HOMA-B and fasting glucose were the tested outcome variables. Gender and BMI were considered as confounders. Fasting insulin and HOMA-IR distributions were skewed because of numerous values below the limit of detection (LOD). This proportion was highly variable, from 5% at age 12 to 39% at age 5. Under such conditions, means and standard errors become unreliable, whereas the median remains an appropriate measure of the distribution of the variable. Therefore, data analysis based on quantile regression (QR) was employed. QR is a non-parametric method based on linear regression that makes no assumptions on the underlying distribution and better suited to deal with data skewness (14). Assuming that the number of values below the LOD did not exceed 40% at any single visit, the median (i.e. 50th percentile) was considered as an analog to the mean of each trait. However, in genetic studies of metabolic traits, genetic variants may display quantile-specific effects (15-17). Quantile regression was thus performed on the 75th percentile for each trait to identify common genetic variants associated with high values of glycemic and insulin traits. Longitudinal association analysis testing association between candidate SNPs and traits trajectory was conducted using mixed-effect QR as implemented in *lqmm* R package (18). SNPs were encoded assuming an additive effect according to the risk allele definition available from PhenoScanner. In addition to the SNP main effect, the SNP x age interaction was also tested. For each trait, multiple testing was controlled by computing the false discovery rate (FDR). Since our analysis was based on candidate SNPs already identified for related glucose / insulin traits, we applied a relaxed 20% FDR significance cutoff. Additional Genetic risk Score were computed and reported in Supplementary materials and Supplementary Tables 4 to 9. Additional longitudinal analyses were performed using the same methodology and criteria to test

associations between SNPs with DHEAS, androstenedione, free testosterone, OHP-17, SHBG, and IGF-1. Association analysis with 50th and 75th percentiles of the hormone distributions was performed only for SNPs significantly associated to 50th and 75th percentile insulin / glucose related traits, respectively.

Role of the funding source

The genotyping and genetic analyses, data interpretation and the writing of the manuscript was funded and performed in collaboration with Nestlé Research.

Results

The population characteristics are reported for the 12-year period in Supplementary Table 10, and Figure 1. The glycemic and insulin traits followed the same trajectories in males and females. However, females had higher HOMA-B levels ($p = 0.02$ for a gender main effect, Figure 1). Fasting glucose increased from age 5 to reach a plateau between ages 13 and 15 years, and then tended to decrease. Trajectories were very similar for insulin and HOMA indices, with decreases from ages 5 to 7, and increases until 12 years in females and 14 in males, after which a plateau was reached, followed by a fall.

To examine how genetic variants, puberty, and weight gain interact to influence insulin action and blood glucose levels, the primary analysis studied the genetic associations with the 50^{th} percentile traits for fasting glucose, fasting insulin, HOMA-IR and HOMA-B across the 12-year period (Table 1, Supplementary Tables 11-14). Assuming a 20% FDR cutoff, 8 SNPs were associated with one or more traits (Table 1), and two of them interacted with age (Supplementary Table 15).

The SNPs rs780094 and rs4457053, two intronic genetic variants of *GCKR* and *ZBED3* respectively, showed a positive main effect on fasting glucose. The first variant showed an age-dependent interaction with the glucose trait (interaction test p = 9.6e-4, adjusted p= 0.11).

HOMA-B carried most of the replicated genetic signal with six negatively associated common variants: rs11257655 located within a DNase I hypersensitive sites in the regulatory region of *CDC12*; rs12779790 within the *CDC123* – *CAMK1D* region; rs1111875 within the *HHEX* – *IDE* gene regions, rs7178572, an intronic SNP of *HMG20A*, rs9787485 within the promoter region of *NRG3*; and the intronic SNP rs1535500 of *KCNK16*. Three of these SNPs - rs9787485, rs11257655 and rs12779790 - were also associated with fasting insulin, and rs9787485 SNP with HOMA-IR too. For rs9787485 SNP, homozygous carriers of the T allele tended to present a different time course for insulin and HOMA indices when compared to allele C carriers (Supplementary Figure 1).

Since elevated fasting glucose and insulin are associated with impaired fasting glucose or T2D, a secondary analysis was performed using quantile regression on the 75th percentile for each trait to identify common genetic variants associated with high values of glycemic and insulin traits (Table 2, Supplementary Tables 16-19) . All but rs1535500 SNP replicated an association with one or more 50th percentile for insulin or glucose traits (Table 2, Supplementary Table 20). Altogether, association signal was extracted for 50 SNPs in or close to 36 genes. The analysis reveals an additional set of SNPs targeting 15 genes likely involved in regulation of high level fasting glucose. For insulin associated traits, no new SNPs were identified.

Puberty is a time during which rapid and dynamic changes occur in hormonal regulations due to increase in sex hormones, that induce transient changes in insulin sensitivity and insulin secretion. As a secondary exploratory analysis, we explored if insulin action may be influenced by genetic

variations in the growth hormone-IGF-1 axis, adrenal and sex steroid activity. Genetic associations with 17-OHP, DHEAS, androstenedione, free testosterone, SHBG, and IGF-1were investigated over the 12-year period only for SNPs previously associated insulin / glucose related traits. The analysis was first conducted usingthe 50^{th} percentile traits for each endocrine parameter(Supplementary Tables 21-26),. As previously, the analysis was performed as well on the 75^{th} percentile endocrine traits (Supplementary Tables 27-32). Of note is the association of the genetic variant rs1111875 within the HHEX – IDE genes region with concentrations of SHBG (main effect test p = 0.0059; FDR=0.047), IGF-1 (age interaction test p = 0.013; FDR=0.106) and 17-OHP (main effect test p = 0.0085; FDR=0.068; age interaction test p = 0.0081, FDR = 0.065), suggesting the genotype-related HOMA-B trajectory is associated with more profound physiological changes through pubertal development.

Discussion

This novel work on the EarlyBird cohort has demonstrated that some genetic risk markers associated with increased risks of T2D also influence the trajectory of insulin and glucose during childhood in children independently of BMI. We report how longitudinal interaction analysis of SNPs and metabolic phenotypes of children has the potential to shed further light on the molecular disturbances associated with insulin resistance in childhood. Some of these genetic risk markers may be particularly relevant for the early identification of children at risk of prediabetes, independently of obesity.

Genome-wide association studies have identified SNPs that are associated with tissue-specific insulin resistance, β-cell dysfunction, or both, making individuals with these variants more prone to the adverse metabolic effects of obesity and type 2 diabetes (19; 20). Yet, due to the difficulty in recruiting large children's cohorts, most genome-wide association studies have been undertaken in

adults. Information about the genetics of prediabetes in children and adolescents remains scarce and mainly limited to cross-sectional studies on children who were already obese (21; 22).

The present analysis on children from the EarlyBird cohort is a significant advance because it sheds light upon changes in insulin secretion, insulin action and glycaemia over time. We found that SNPs associated with adult diabetes susceptibility, not only replicated their association with insulin and glycemic traits in these children independently of BMI, but also showed age-specific interactions with some of these traits (50th percentile traits analysis), including rs780094 (*GCKR*), rs4457053 (*ZBED3)*, rs11257655 (*CDC123)*, rs12779790 (*CDC123*, *CAMK1D*), rs1111875 (*HHEX*), rs7178572 (*HMG20A*), rs9787485 (*NRG3*), and rs1535500 (*KCNK16*). A greater number of common genetic variants were associated with variation of high level of fasting glucose distribution (75^{th} percentile). This observation is consistent with the understanding that elevated fasting plasma glucose is a risk factor for T2D, and previous evidence that selected SNPs were associated with T2D in adult populations (23).

Two intronic genetic variants of *GCKR* and *ZBED3* were associated with different trajectories of fasting glucose from an early age throughout childhood for both the median and higher (75^{th} percentile) trait distributions. The SNP in *ZEBD3* may contribute to the risk of T2D through elevated WNT activity (24). The WNT pathway plays involves pancreas β–cell genesis, GLP-1-mediated proliferation, and synthesis of GLP-1 (24). Moreover, the activity of this pathway has been shown to be modulated by short chain fatty acid produced during digestion of dietary fibers, and thus partially mediating their anti-diabetogenic effects (24). Since the SNP in *ZBED3* was shown previously to influence the effect of fiber intake on the incidence of T2D (24), it may be a relevant genetic marker for children who would benefit from optimized dietary fiber intake. Glucokinase regulatory protein (GCKR) regulates the activity of glucokinase in the liver and the pancreatic β-cells, with a pivotal role in glucose-stimulated insulin release and systemic glucose

homeostasis (25). The intragenic SNP of *GCKR* contributes to the risk of T2D and dyslipidemia in different populations, and may be associated with lower fasting blood glucose levels in adults (25). We found that the *GCKR* genotype was negatively associated with an age-dependent course of glucose, yet showed a positive association with blood glucose at age 5. Glucose metabolism of β-cells also regulates the adaptive response to fuel loads (26), and it has been shown that children exhibit increased carbohydrate oxidation during pubertal development (27). Therefore, understanding of the role of β-cells in regulating fuel homeostasis during growth and development in healthy children is likely to be relevant for diabetes risk management (26).

Obese children and adolescents developing (pre-) diabetes generally have a higher genetic predisposition related to gene variants modulating the early, dynamic phase of insulin secretion (20). In contrast with previous studies, we identified that defects in β-cell function distinguished children who developed prediabetes in the EarlyBird cohort, independently of BMI (6). Our analysis has found that among common variants associated with diabetes susceptibility, HOMA-B carried most of the replicated genetic associations, independently of BMI. Some of the replicated genetic variants, namely SNPs in the *CDC123, HHEX-IDE* and *KCNK16* loci, showed a negative association with HOMA-B in the Earlybird cohort. Our observations are in agreement with reports in various populations of their associations with reduced β-cell function and insulin secretion (20; 28; 29), and the proposition that SNPs in the *CDC123 loci* may represent a proxy for β-cell mass (28). Therefore, these genetic variants, previously identified as risk markers, also influence the normal course of β-cell function during growth and development of healthy children.

We explored whether genotypes also influence endocrine traits, and therefore the maturation of other biological processes in childhood. We found some evidence for this in the interactions of the SNP in *HHEX-IDE* loci with the trajectories of SHBG, IGF-1 and 17-OHP. These findings support

the biologically plausible idea that insulin action is also influenced by genetic variations in the growth hormone-IGF-1 axis, adrenal and sex steroid activity. Of particular note is the negative association with both HOMA-B and SHBG. SHBG is a glycoprotein that transports sex steroids in the circulation and regulates their access to target cells, and for which low levels have been linked to diabetes and early puberty (30). Since the HHEX-IDE SNP relates to insulin secretion capacity in early life (31), our observations suggest effects beyond fetal development into pubertal growth and development.

This study has several limitations, including limited sample size and ethnic homogeneity of the EarlyBird cohort. However, the inclusion of healthy children who were of normal weight and insulin sensitivity reduces the variance of insulin and HOMA indices in the study population. A significant strength of the EarlyBird study is that it is a truly longitudinal study of a cohort of healthy children, with detailed annual phenotypic measurements from age 5 to 16 years. Importantly, the majority of the children were of normal weight and the insulin resistance observed was within the physiological range. Therefore, the EarlyBird study is not confounded by the effects of pre-existing obesity and insulin resistance, rather it provides insights into the beginnings of these conditions.

In conclusion, we demonstrated that SNPs previously associated with diabetes in adults also influence the course of glycemic and insulin traits during childhood independently of BMI. Our study has potential clinical applications, since β-cell dysfunction is an early event in the pathogenesis of diabetes. Further weight gain and demand for insulin will aggravate the progression from prediabetes to overt diabetes (32). Eight of these genetic risk markers may be particularly relevant for future studies aiming at early identification of children at risk of prediabetes, before puberty and prior to the development of obesity. Individuals at risk potentially could be offered an

early lifestyle modification program to reduce the risks of progression towards diabetes. An important aim for future research will be to understand mechanisms of how genetic factors, puberty and weight gain interact in the pathogenesis of pre-diabetes. This knowledge could allow the development of disease prevention strategies in children, using personalized lifestyle and dietary interventions.

Acknowledgements

We acknowledge the life and work of our former colleague Terence Wilkin (1945-2017), Professor of Endocrinology and Metabolism, whose vision and original thinking led to the creation of the EarlyBird Study and the establishment of the collaboration that made possible the genetic study reported here. We thank Ondine Walter for biobanking, sample handling and preparation at Nestlé, and for support for compliance with the Human Research Act. We thank Kei Sakamoto, PhD and Ralf Heine, MD for critical review of the manuscript.

Contribution statement

JHager and FPM designed the study.

AJ, JM, AC, SM were involved in the acquisition of the data.

JC, JM, JHager, FPM, JHosking, JP were involved in the analysis and interpretation of the data.

JC, JHager, JHosking, JP, FPM drafted the manuscript and all co-authors approved the final version.

JP is guarantor of the work.

Duality of interest

JHosking, JP and AJ are employees of Plymouth University Peninsula School of Medicine and Dentistry.

JC, JM, AC, SM, JHager, and FPM are or were employees of the Nestle Group at the time the work has been performed.

JHosking and AJ have received funding from the Nestle Group.

The authors have no other dualities of interest to declare.

Funding

The EarlyBird study was supported by Bright Future Trust, The Kirby Laing Foundation, Peninsula Medical Foundation, Diabetes UK, and the EarlyBird Diabetes Trust. The genetic analysis reported in this paper was funded by Nestlé Research.

Data sharing statement

Data may be available upon request to Francois-Pierre Martin and Jonathan Pinkney, subject in particular, to ethical and privacy considerations.

References

1. Organization WH: Global report on diabetes. 2016;
2. Narayan KM, Boyle JP, Thompson TJ, Sorensen SW, Williamson DF: Lifetime risk for diabetes mellitus in the United States. JAMA 2003;290:1884-1890
3. Mayer-Davis EJ, Dabelea D, Lawrence JM: Incidence Trends of Type 1 and Type 2 Diabetes among Youths, 2002-2012. The New England journal of medicine 2017;377:301
4. Wilkin TJ: The accelerator hypothesis: weight gain as the missing link between Type I and Type II diabetes. Diabetologia 2001;44:914-922
5. Weiss R, Santoro N, Giannini C, Galderisi A, Umano GR, Caprio S: Prediabetes in youth - mechanisms and biomarkers. Lancet Child Adolesc Health 2017;1:240-248
6. Hosking J, Metcalf BS, Jeffery AN, Streeter AJ, Voss LD, Wilkin TJ: Evidence of early beta-cell deficiency among children who show impaired fasting glucose: 10-yr cohort study (EarlyBird 56). Pediatric diabetes 2013;14:481-489
7. Li C, Ford ES, Zhao G, Mokdad AH: Prevalence of pre-diabetes and its association with clustering of cardiometabolic risk factors and hyperinsulinemia among U.S. adolescents: National Health and Nutrition Examination Survey 2005-2006. Diabetes Care 2009;32:342-347
8. Lauria M, Persico M, Dordevic N, Cominetti O, Matone A, Hosking J, Jeffery A, Pinkney J, Da Silva L, Priami C, Montoliu I, Martin FP: Consensus Clustering of temporal profiles for the identification of metabolic markers of pre-diabetes in childhood (EarlyBird 73). Scientific reports 2018;8:1393
9. Voss LD, Kirkby J, Metcalf BS, Jeffery AN, O'Riordan C, Murphy MJ, Wilkin TJ: Preventable factors in childhood that lead to insulin resistance, diabetes mellitus and the metabolic syndrome: the EarlyBird diabetes study 1. J Pediatr Endocrinol Metab 2003;16:1211-1224
10. Cole TJ, Freeman JV, Preece MA: Body mass index reference curves for the UK, 1990. Archives of disease in childhood 1995;73:25-29
11. Gungor N, Saad R, Janosky J, Arslanian S: Validation of surrogate estimates of insulin sensitivity and insulin secretion in children and adolescents. The Journal of pediatrics 2004;144:47-55
12. Vermeulen A, Verdonck L, Kaufman JM: A critical evaluation of simple methods for the estimation of free testosterone in serum. The Journal of clinical endocrinology and metabolism 1999;84:3666-3672
13. MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, Junkins H, McMahon A, Milano A, Morales J, Pendlington ZM, Welter D, Burdett T, Hindorff L, Flicek P, Cunningham F, Parkinson H: The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic acids research 2017;45:D896-D901
14. Inference for Quantile Regression. In *Quantile Regression* Koenker R, Ed. Cambridge, Cambridge University Press, 2005, p. 68-115
15. Williams PT: Quantile-specific penetrance of genes affecting lipoproteins, adiposity and height. PloS one 2012;7:e28764
16. White MJ, Kodaman NM, Harder RH, Asselbergs FW, Vaughan DE, Brown NJ, Moore JH, Williams SM: Genetics of Plasminogen Activator Inhibitor-1 (PAI-1) in a Ghanaian Population. PloS one 2015;10:e0136379
17. Briollais L, Durrieu G: Application of quantile regression to recent genetic and -omic studies. Human genetics 2014;133:951-966
18. Geraci M: Linear Quantile Mixed Models: The lqmm Package for Laplace Quantile Regression. 2014 2014;57:29
19. Goodarzi MO: Genetics of obesity: what genetic association studies have taught us about the biology of obesity and its complications. Lancet Diabetes Endocrinol 2018;6:223-236
20. Giannini C, Dalla Man C, Groop L, Cobelli C, Zhao H, Shaw MM, Duran E, Pierpont B, Bale AE, Caprio S, Santoro N: Co-occurrence of risk alleles in or near genes modulating insulin secretion predisposes obese youth to prediabetes. Diabetes Care 2014;37:475-482

21. Barker A, Sharp SJ, Timpson NJ, Bouatia-Naji N, Warrington NM, Kanoni S, Beilin LJ, Brage S, Deloukas P, Evans DM, Grontved A, Hassanali N, Lawlor DA, Lecoeur C, Loos RJ, Lye SJ, McCarthy MI, Mori TA, Ndiaye NC, Newnham JP, Ntalla I, Pennell CE, St Pourcain B, Prokopenko I, Ring SM, Sattar N, Visvikis-Siest S, Dedoussis GV, Palmer LJ, Froguel P, Smith GD, Ekelund U, Wareham NJ, Langenberg C: Association of genetic Loci with glucose levels in childhood and adolescence: a meta-analysis of over 6,000 children. Diabetes 2011;60:1805-1812
22. Kelliny C, Ekelund U, Andersen LB, Brage S, Loos RJ, Wareham NJ, Langenberg C: Common genetic determinants of glucose homeostasis in healthy children: the European Youth Heart Study. Diabetes 2009;58:2939-2945
23. Nguyen QM, Srinivasan SR, Xu JH, Chen W, Berenson GS: Fasting plasma glucose levels within the normoglycemic range in childhood as a predictor of prediabetes and type 2 diabetes in adulthood: the Bogalusa Heart Study. Arch Pediatr Adolesc Med 2010;164:124-128
24. Hindy G, Mollet IG, Rukh G, Ericson U, Orho-Melander M: Several type 2 diabetes-associated variants in genes annotated to WNT signaling interact with dietary fiber in relation to incidence of type 2 diabetes. Genes Nutr 2016;11:6
25. Bi M, Kao WH, Boerwinkle E, Hoogeveen RC, Rasmussen-Torvik LJ, Astor BC, North KE, Coresh J, Kottgen A: Association of rs780094 in GCKR with metabolic traits and incident diabetes and cardiovascular disease: the ARIC Study. PloS one 2010;5:e11690
26. Matschinsky FM: Banting Lecture 1995. A lesson in metabolic regulation inspired by the glucokinase glucose sensor paradigm. Diabetes 1996;45:223-241
27. Caprio S, Cline G, Boulware S, Permanente C, Shulman GI, Sherwin RS, Tamborlane WV: Effects of puberty and diabetes on metabolism of insulin-sensitive fuels. Am J Physiol 1994;266:E885-891
28. Simonis-Bik AM, Nijpels G, van Haeften TW, Houwing-Duistermaat JJ, Boomsma DI, Reiling E, van Hove EC, Diamant M, Kramer MH, Heine RJ, Maassen JA, Slagboom PE, Willemsen G, Dekker JM, Eekhoff EM, de Geus EJ, t Hart LM: Gene variants in the novel type 2 diabetes loci CDC123/CAMK1D, THADA, ADAMTS9, BCL11A, and MTNR1B affect different aspects of pancreatic beta-cell function. Diabetes 2010;59:293-301
29. Varshney A, Scott LJ, Welch RP, Erdos MR, Chines PS, Narisu N, Albanus RD, Orchard P, Wolford BN, Kursawe R, Vadlamudi S, Cannon ME, Didion JP, Hensley J, Kirilusha A, Program NCS, Bonnycastle LL, Taylor DL, Watanabe R, Mohlke KL, Boehnke M, Collins FS, Parker SC, Stitzel ML: Genetic regulatory signatures underlying islet gene expression and type 2 diabetes. Proc Natl Acad Sci U S A 2017;114:2301-2306
30. Aydin B, Winters SJ: Sex Hormone-Binding Globulin in Children and Adolescents. J Clin Res Pediatr Endocrinol 2016;8:1-12
31. Winkler C, Illig T, Koczwara K, Bonifacio E, Ziegler AG: HHEX-IDE polymorphism is associated with low birth weight in offspring with a family history of type 1 diabetes. The Journal of clinical endocrinology and metabolism 2009;94:4113-4115
32. Chen C, Cohrs CM, Stertmann J, Bozsak R, Speier S: Human beta cell mass and function in diabetes: Recent advances in knowledge and technologies to understand disease pathogenesis. Mol Metab 2017;6:943-957

Figure 1. Box-plot distributions of glycemic and insulin traits from 5 to 16 years old in males (red) and females (blue).

Table 1: Results from mixed-linear quantile regression for main effects of SNPs on 50th percentile distribution of insulin and glycemic traits trajectory from 5 to 16 years old.

						Fasting glucose		**Fasting insulin**		**HOMA-B**		**HOMA-IR**	
SNP	Risk allele	Freq	Region	Reported gene[a]	Initial GWAS phenotype	Coef	P	Coef	P	Coef	P	Coef	P
rs11257655	T	0.26	10p13	CDC123	Type 2 diabetes	**0.028**	**0.54**	**-1.26**	**0.0022***	**-16.5808**	**0.00043****	**-0.16**	**0.0055**
rs12779790	G	0.28	10p13	CDC123, CAMK1D	Type 2 diabetes	**0.031**	**0.51**	**-1.26**	**0.0027***	**-15.53**	**0.0011****	**-0.16**	**0.0083**
rs1111875	A	0.4	10q23.3	IDE	Type 2 diabetes	**0.062**	**0.093**	**-0.36**	**0.24**	**-8.75**	**0.0095***	**-0.048**	**0.2**
rs7178572	A	0.23	15q24.3	HMG20A	Type 2 diabetes	**0.026**	**0.46**	**-0.61**	**0.056**	**-9.22**	**0.0091***	**-0.077**	**0.091**
rs9787485	T	0.17	10q23.1	NRG3	HOMA-IR	**-0.018**	**0.62**	**-1.33**	**0.0029***	**-15.054**	**0.0027****	**-0.21**	**0.0015***
rs780094	A	0.34	2p23.3	GCKR	Type 2 diabetes	**0.13**	**0.001***	**0.25**	**0.46**	**-4.079**	**0.26**	**0.018**	**0.7**
rs4457053	G	0.32	5q13.3	ZBED3	Type 2 diabetes	**0.13**	**0.0022***	**-0.07**	**0.85**	**-4.32**	**0.32**	**-0.0034**	**0.95**
rs1535500	G	0.43	6p21.2	KCNK16	Type 2 diabetes	**-0.013**	**0.72**	**-0.93**	**0.018**	**-14.001**	**0.0018****	**-0.13**	**0.016**

Note Bene: Results are provided for SNPs with significant (FDR cutoff set to 20%) main test with at least one of the four traits. Coef: Coefficient indicating the directions of the associations between the SNPs and the glycemic or insulin trait. Test pvalue with FDR cutoff set to * 20%; ** 10%; [a]Genes reported in GWAS catalog.

Table 2: Results from mixed-linear quantile regression for main effects of SNPs on 75th percentile distribution of insulin and glycemic traits trajectory from 5 to 16 years old.

						Fasting glucose		**Fasting insulin**		**HOMA-B**		**HOMA-IR**	
SNP	Risk allele	Freq	Region	Reported gene[a]	Initial GWAS phenotype	Coef	P	Coef	P	Coef	P	Coef	P
rs702634	G	0.37	5q11.2	ARL15	Type 2 diabetes	0.092	0.02**	-0.12	0.72	-2.51	0.50	-0.031	0.53
rs243021	T	0.46	2p16.1	BCL11A	Type 2 diabetes	0.12	0.0059***	0.24	0.47	3.97	0.35	0.020	0.66
rs243088	A	0.48	2p16.1	BCL11A	Type 2 diabetes	0.087	0.038*	0.38	0.19	4.96	0.17	0.043	0.30
rs2812533	T	0.12	10q22.1	C10orf35	Type 2 diabetes	0.14	0.0033***	1.03	0.049	10.98	0.059	0.146	0.052
rs11257655	T	0.26	10p13	CDC123	Type 2 diabetes	0.061	0.29	-1.19	0.0034	-14.96	0.000891*	-0.16	0.0067
rs12779790	G	0.28	10p13	CDC123	Type 2 diabetes	0.082	0.24	-1.18	0.0051	-13.93	0.002434*	-0.16	0.011
rs10440833	A	0.26	6p22.3	CDKAL1	Type 2 diabetes	0.099	0.018**	0.73	0.040	5.33	0.22	0.104	0.041
rs4712523	G	0.33	6p22.3	CDKAL1	Type 2 diabetes	0.11	0.037*	0.66	0.081	5.56	0.22	0.086	0.11
rs6931514	G	0.26	6p22.3	CDKAL1	Type 2 diabetes	0.099	0.018**	0.73	0.040	5.33	0.22	0.104	0.041
rs7756992	G	0.27	6p22.3	CDKAL1	Type 2 diabetes	0.099	0.018**	0.73	0.040	5.33	0.22	0.104	0.041
rs7766070	A	0.27	6p22.3	CDKAL1	Type 2 diabetes	0.15	0.001***	0.75	0.046	5.49	0.23	0.105	0.056
rs7018475	G	0.2	9p21.3	CDKN2B	Type 2 diabetes	0.17	3.4e-05***	-0.42	0.35	-6.42	0.21	-0.056	0.37
rs9841287	G	0.16	3p26.3	CHL1	HOMA-IR	0.15	0.010***	0.47	0.40	4.99	0.42	0.045	0.58
rs7607980	C	0.13	2q24.3	COBLL1	Type 2 diabetes	0.082	0.20	1.19	0.036	14.37	0.042	0.161	0.044
rs2284219	A	0.34	7p14.3	CRHR2	Type 2 diabetes	0.093	0.0099***	-0.050	0.88	0.11	0.98	-0.009	0.85
rs1153188	A	0.22	12q13.2	DCD	Type 2 diabetes	0.085	0.087	1.05	0.014	8.90	0.088	0.127	0.035
rs11642841	A	0.35	16q12.2	FTO	Type 2 diabetes	0.14	0.0053***	-0.15	0.63	-1.47	0.64	-0.023	0.62
rs8050136	A	0.4	16q12.2	FTO	Type 2 diabetes	0.21	0.00021***	0.035	0.91	-1.65	0.63	-0.0071	0.87
rs9936385	C	0.4	16q12.2	FTO	Type 2 diabetes	0.21	0.00021***	0.035	0.91	-1.65	0.63	-0.0071	0.87

rs9939609	A	0.41	16q12.2	FTO	Type 2 diabetes	0.10	0.01***	0.045	0.89	-1.65	0.66	-0.0050	0.92
rs780094	A	0.34	2p23.3	GCKR	HOMA-IR	0.21	2.7e-05***	0.37	0.22	-0.56	0.86	0.043	0.34
rs3923113	G	0.35	2q24.3	GRB14	Type 2 diabetes	0.032	0.48	0.69	0.052	11.66	0.004609*	0.087	0.095
rs1111875	A	0.4	10q23.33	HHEX, IDE	Type 2 diabetes	0.17	0.00024***	-0.26	0.47	-5.78	0.14	-0.042	0.44
rs5015480	T	0.4	10q23.33	HHEX, IDE	Type 2 diabetes	0.11	0.0049***	-0.24	0.54	-5.61	0.19	-0.044	0.46
rs7178572	A	0.23	15q24.3	HMG20A	Type 2 diabetes	0.063	0.096	-0.50	0.061	-7.29	0.017	-0.070	0.075
rs1531343	C	0.07	12q14.3	HMGA2	Type 2 diabetes	0.23	0.00043***	-0.63	0.38	-12.28	0.11	-0.095	0.36
rs2261181	T	0.09	12q14.3	HMGA2	Type 2 diabetes	0.13	0.028*	-0.12	0.86	-7.30	0.33	-0.025	0.81
rs4430796	G	0.47	17q12	HNF1B	Type 2 diabetes	0.11	0.0021***	0.11	0.73	1.31	0.72	0.0018	0.97
rs2943640	A	0.31	2q36.3	IRS1	Type 2 diabetes	0.051	0.25	0.50	0.17	5.23	0.19	0.060	0.25
rs849134	G	0.5	7p15.1	JAZF1	Type 2 diabetes	0.091	0.031*	0.41	0.18	3.39	0.35	0.056	0.20
rs849135	A	0.5	7p15.1	JAZF1	Type 2 diabetes	0.099	0.0087***	0.42	0.18	3.60	0.33	0.058	0.20
rs864745	G	0.5	7p15.1	JAZF1	Type 2 diabetes	0.080	0.036*	0.36	0.23	3.23	0.36	0.048	0.25
rs5215	C	0.36	11p15.1	KCNJ11	Type 2 diabetes	0.091	0.032*	0.25	0.45	2.86	0.48	0.031	0.54
rs972283	A	0.49	7q32.3	KLF14	Type 2 diabetes	0.095	0.0094***	0.061	0.88	3.77	0.36	0.0030	0.96
rs10842994	T	0.14	12p11.22	KLHDC5	Type 2 diabetes	0.19	0.0011***	0.40	0.44	6.08	0.28	0.061	0.44
rs2943641	T	0.31	2q36.3	IRS1	Type 2 diabetes	0.056	0.26	0.46	0.20	4.95	0.22	0.056	0.28
rs9787485	T	0.17	10q23.1	NRG3	HOMA-IR	-0.0037	0.93	-1.29	0.0027	-13.38	0.005688*	-0.20	0.0012*
rs12970134	A	0.19	18q21.32	MC4R	Type 2 diabetes	0.10	0.035*	0.27	0.45	3.10	0.48	0.019	0.71
rs1387153	T	0.32	11q14.3	MTNR1B	Type 2 diabetes	0.093	0.0059***	0.57	0.12	5.28	0.22	0.092	0.10
rs8182584	T	0.39	19q13.11	PEPD	Fasting Insulin	0.12	0.0032***	0.16	0.61	0.84	0.80	0.012	0.78
rs13081389	G	0.12	3p25.2	PPARG	Type 2 diabetes	0.14	0.021**	0.99	0.15	7.18	0.39	0.138	0.15
rs1801282	G	0.17	3p25.2	PPARG	Type 2 diabetes	0.13	0.028*	0.80	0.12	8.99	0.17	0.107	0.13
rs12899811	G	0.3	15q26.1	PRC1	Type 2 diabetes	0.078	0.057*	0.76	0.048	10.20	0.024	0.097	0.085
rs8042680	A	0.29	15q26.1	PRC1	Type 2 diabetes	0.14	0.0026***	0.63	0.12	10.38	0.023	0.079	0.19

rs1359790	T	0.23	13q31.1	SPRY2	Type 2 diabetes	-0.11	0.0074***	-0.64	0.088	-3.78	0.37	-0.097	0.074
rs13273088	G	0.23	8q13.2	SULF1	Fasting Insulin	0.11	0.024**	-0.29	0.48	-4.92	0.18	-0.028	0.63
rs17791513	G	0.09	9q21.31	TLE4	Type 2 diabetes	0.17	0.049*	0.84	0.12	7.64	0.23	0.143	0.068
rs4760790	A	0.27	12q21.1	LGR5	Type 2 diabetes	0.087	0.016**	0.47	0.17	2.36	0.54	0.063	0.20
rs4457053	G	0.32	5q13.3	ZBED3	Type 2 diabetes	0.17	0.00013***	-0.083	0.82	-2.28	0.56	0.029	0.57
rs12571751	G	0.49	10q22.3	ZMIZ1	Type 2 diabetes	0.13	0.00024***	0.039	0.90	-0.53	0.87	-0.0034	0.94

Note Bene: Results are provided for SNPs with significant (FDR cutoff set to 20%) main test with at least one of the four traits. Coef: Coefficient indicating the directions of the associations between the SNPs and the glycemic or insulin trait. Test pvalue with FDR cutoff set to *20%; ** 10%; *** 5%; [a]Genes reported in GWAS catalog.