

Article

The Relationship of Maternal Gestational Mass Spectrometry-Derived Metabolites with Offspring Congenital Heart Disease: Results from Multivariable and Mendelian Randomization Analyses

Kurt Taylor ^{1,2,*}, Nancy McBride ^{1,2}, Jian Zhao ^{1,3,4}, Sam Oddie ⁵, Razaq Azad ⁶, John Wright ⁶, Ole A. Andreassen ^{7,8}, Isobel D. Stewart ⁹, Claudia Langenberg ^{9,10,11}, Maria Christine Magnus ^{1,2,12}, Maria Carolina Borges ^{1,2}, Massimo Caputo ^{13,14} and Deborah A. Lawlor ^{1,2,13}

- ¹ Population Health Science, Bristol Medical School, Bristol BS8 2PS, UK; nancy.mcbride@bristol.ac.uk (N.M.); jian.zhao@bristol.ac.uk (J.Z.); mariachristine.magnus@fhi.no (M.C.M.); m.c.borges@bristol.ac.uk (M.C.B.); d.a.lawlor@bristol.ac.uk (D.A.L.)
 - ² MRC Integrative Epidemiology Unit, University of Bristol, Bristol BS8 2BN, UK
 - ³ The Ministry of Education and Shanghai Key Laboratory of Children's Environmental Health, Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200092, China
 - ⁴ Department of Maternal and Child Health, School of Public Health, Shanghai Jiao Tong University, Shanghai 200240, China
 - ⁵ The Hull York Medical School, University of York, Heslington YO10 5DD, UK; sam.oddie@bthft.nhs.uk
 - ⁶ Bradford Institute for Health Research, Bradford Teaching Hospitals National Health Service Foundation Trust, Bradford BD9 6RJ, UK; rafaq.azad@bthft.nhs.uk (R.A.); john.wright@bthft.nhs.uk (J.W.)
 - ⁷ NORMENT Centre, Institute of Clinical Medicine, Division of Mental Health and Addiction, Oslo University Hospital, University of Oslo, 0315 Oslo, Norway; ole.andreassen@medisin.uio.no
 - ⁸ KG Jebsen Centre for Neurodevelopmental Disorders, Institute of Clinical Medicine, Oslo University Hospital, 0424 Oslo, Norway
 - ⁹ MRC Epidemiology Unit, University of Cambridge, Cambridge CB2 0SL, UK; isobel.stewart@mrc-epid.cam.ac.uk (I.D.S.); claudia.langenberg@mrc-epid.cam.ac.uk (C.L.)
 - ¹⁰ Health Data Research UK Cambridge, Wellcome Genome Campus and University of Cambridge, Cambridge CB10 1SA, UK
 - ¹¹ Computational Medicine, Berlin Institute of Health (BIH), Charité University Medicine, 10178 Berlin, Germany
 - ¹² Centre for Fertility and Health, Norwegian Institute of Public Health, 0473 Oslo, Norway
 - ¹³ National Institute for Health Research Bristol Biomedical Centre, University Hospitals Bristol NHS Foundation Trust, University of Bristol, Bristol BS8 2BN, UK; m.caputo@bristol.ac.uk
 - ¹⁴ Translational Science, Bristol Medical School, Bristol BS2 8HW, UK
- * Correspondence: kurt.taylor@bristol.ac.uk



Citation: Taylor, K.; McBride, N.; Zhao, J.; Oddie, S.; Azad, R.; Wright, J.; Andreassen, O.A.; Stewart, I.D.; Langenberg, C.; Magnus, M.C.; et al. The Relationship of Maternal Gestational Mass Spectrometry-Derived Metabolites with Offspring Congenital Heart Disease: Results from Multivariable and Mendelian Randomization Analyses. *J. Cardiovasc. Dev. Dis.* **2022**, *9*, 237. <https://doi.org/10.3390/jcdd9080237>

Academic Editor: Maurice Van den Hoff

Received: 12 May 2022

Accepted: 22 July 2022

Published: 27 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Background: It is plausible that maternal pregnancy metabolism influences the risk of offspring congenital heart disease (CHD). We sought to explore this through a systematic approach using different methods and data. **Methods:** We undertook multivariable logistic regression of the odds of CHD for 923 mass spectrometry (MS)-derived metabolites in a sub-sample of a UK birth cohort (Born in Bradford (BiB); N = 2605, 46 CHD cases). We considered metabolites reaching a *p*-value threshold <0.05 to be suggestively associated with CHD. We sought validation of our findings, by repeating the multivariable regression analysis within the BiB cohort for any suggestively associated metabolite that was measured by nuclear magnetic resonance (NMR) or clinical chemistry (N = 7296, 87 CHD cases), and by using genetic risk scores (GRS: weighted genetic risk scores of single nucleotide polymorphisms (SNPs) that were associated with any suggestive metabolite) in Mendelian randomization (MR) analyses. The MR analyses were performed in BiB and two additional European birth cohorts (N = 38,662, 319 CHD cases). **Results:** In the main multivariable analyses, we identified 44 metabolites suggestively associated with CHD, including those from the following super pathways: amino acids, lipids, co-factors and vitamins, xenobiotics, nucleotides, energy, and several unknown molecules. Of these 44, isoleucine and leucine were available in the larger BiB cohort (NMR), and for these the results were validated. The MR analyses were possible for 27/44 metabolites and for 11 there was consistency with the multivariable regression results. **Conclusions:** In summary, we have

used complimentary data sources and statistical techniques to construct layers of evidence. We found that pregnancy amino acid metabolism, androgenic steroid lipids, and levels of succinylcarnitine could be important contributing factors for CHD.

Keywords: metabolites; congenital heart disease; Born in Bradford; ALSPAC; MoBa

1. Introduction

Congenital heart diseases (CHDs) are the most common congenital anomaly affecting approximately 6–8 per 1000 live births and 10% of stillbirths. They are the leading cause of death from congenital anomalies [1]. Approximately 20% of CHD cases can be attributed to known chromosomal anomalies, gene disorders or teratogens [2]. The causes of the remaining cases are unknown. Identifying causes of CHDs is important for improving aetiological understanding and developing potential targets for intervention.

Metabolomics technologies have enabled the quantification of a large number of metabolites in a biological sample. Metabolites are small-molecule intermediates and products of metabolism. The metabolome, the complete set of metabolites in biological tissues/fluids, is influenced by both genotype and the environment, and dynamically responds to environmental influences. Analyses of maternal metabolomic profiles could identify causal mechanisms leading to CHDs [3]. Because the metabolome reflects interactions of the genomic, environmental (e.g., air pollution), behavioural (e.g., smoking) and pathophysiological states (e.g., body composition), examining associations of it with CHDs could help elucidate modifiable upstream risk factors and/or potential molecular targets for intervention to prevent CHDs.

Studies have explored maternal biomarkers and found that the offspring of women with low 25-hydroxyvitamin D (<50 nmol/L compared to >75 nmol/L) [4] and dyslipidaemic profiles [5,6] have an increased risk of CHDs. Other work has shown that higher fasting glucose levels and diabetes during pregnancy can increase CHD risk [7–9]; however, these studies focus on single or only a few biomarkers. Exploring the wider metabolome could provide opportunities to improve our understanding of the molecular mechanisms that underpin CHDs [3].

One previous study has explored metabolomics in maternal serum including 27 CHD cases compared with 59 controls [10]. A total of 123 (out of 174) metabolites were associated with CHD in univariable analyses, with 3 of the MS metabolites (ε-hydroxypropionylcarnitine, glutacoylcarnitine, and hydroxytetradecadienylcarnitine) in multivariable analyses showing a high discrimination for a first trimester CHD diagnosis (area under the curve 0.98 (0.94, 0.99)). A further study used an untargeted approach to explore whether mid-pregnancy urinary metabolites could accurately discriminate CHD cases from the controls (N = 70 CHD cases and 70 controls) and found that two principal components separated the cases and the controls, with 4-hydroxybenzeneacetic acid, 5-trimethylsilyloxy-n-valeric acid, hydracrylic acid and propanedioic acid driving their differences [11]. Finally, a study in a Chinese population of metabolomic analyses using maternal amniotic fluid (N = 71 CHD cases) found that two metabolites (uric acid and proline) had higher concentrations in CHD affected pregnancies [12]. In summary, to date there have been few pregnancy metabolomic studies exploring the potential molecular causal mechanisms of CHD. The small number that we have identified focused on discrimination between cases and the controls and had not obtained external validation.

The aim of this study was to gain insights into the potential molecular mechanisms for CHD. To address this aim, we explored associations of maternal metabolites quantified by an untargeted mass spectrometry (MS) platform and the odds of CHD in offspring. We searched for relevant studies within The LifeCycle Project-EU Child Cohort Network [13] to identify any study with detailed, untargeted, maternal gestational metabolomic data and offspring CHD information. We identified only one cohort with relevant data in a

subgroup: the Born in Bradford (BiB) cohort (N = 2605 participants; 46 CHD cases) [14,15]. Recognising that these novel data were potentially underpowered, we sought internal validation of the metabolites suggestively associated with CHD, by repeating the multivariable regression analysis within the BiB cohort for any metabolite that was measured by nuclear magnetic resonance (NMR) or clinical chemistry in larger numbers (N = 7296, 87 CHD cases (including those in the original analysis)). We subsequently searched the MR-PREG consortia studies [16,17] for cohorts with maternal genome-wide data and offspring CHD information that could be used for Mendelian randomization (MR) analyses of the associations of genetic instruments for maternal metabolites. We performed pooled MR analyses across three MR-PREG cohorts meeting our criteria (N = 38,663, 319 CHD cases) for any metabolites that were: (i) suggestively associated with CHD in the BiB cohort ($p < 0.05$) and (ii) had summary data in the most recent metabolomic genome-wide association study (GWAS).

2. Methods

2.1. Study Design and Participants

A schematic overview of the study design is illustrated in Figure 1. We excluded children of multiple births because they differ from single births for congenital anomaly outcomes [18,19]. For the multivariable metabolomic analyses, we used data from the BiB cohort, as this was the only cohort that had measures of a substantial number of metabolites reflecting a wide range of metabolic paths assessed during pregnancy and CHD outcomes [15]. We also explored internal validation of any findings with a p -value < 0.05 within the BiB study where equivalent (or near equivalent) measures to any on the MS platform were available from other sources. BiB is a population-based prospective birth cohort, including 12,453 women across 13,776 pregnancies who were recruited at their oral glucose tolerance test (OGTT) at approximately 26–28 weeks' gestation [14]. The eligible women had an expected delivery date between March 2007 and December 2010. The use of a multivariable p -value threshold of < 0.05 to take associations forward into further validation analyses was appropriate as an initial screen, for a relatively rare outcome, to avoid missing potential causal effects. All the BiB participants included in the Metabolon multivariable regression analyses (N = 2605) (Step 1 (Figure 1)) were also included in the NMR validation analyses (Step 2a (Figure 1)), and 2147 had the relevant genetic data and were, therefore, included in the MR analyses.

To be included in the MR analyses, the studies and participants had to have genome-wide data in the mothers and CHD data in their offspring. Three cohorts contributed to the MR analyses: BiB, the Avon Longitudinal Study of Parents and Children (ALSPAC) and the Norwegian Mother, Father and Child Cohort Study (MoBa). ALSPAC is a UK prospective birth cohort study which was devised to investigate the environmental and genetic factors of health and development [20–22]. Pregnant women resident in Avon, UK with expected dates of delivery 1 April 1991 to 31 December 1992 were invited to take part in the study. The initial number of pregnancies enrolled was 14,541 (and for these at least one questionnaire was returned, or a "Children in Focus" clinic had been attended by 19/07/99). Of those initial pregnancies, there was a total of 14,676 fetuses, resulting in 14,062 live births and 13,988 children who were alive at 1 year of age. MoBa is a population-based pregnancy cohort study conducted by the Norwegian Institute of Public Health [23,24]. The participants were recruited from all over Norway from 1999–2008. The women consented to participation in 41% of the pregnancies. The cohort included approximately 114,500 children, 95,200 mothers and 75,200 fathers. The current study is based on 12 of the quality-assured data files released for research in 2019. The establishment of the MoBa and initial data collection was based on a license from the Norwegian Data Protection Agency and approval from The Regional Committees for Medical and Health Research Ethics. The MoBa cohort is currently regulated by the Norwegian Health Registry Act. The current study was approved by The Regional Committees for Medical and Health Research Ethics (2018/1256).

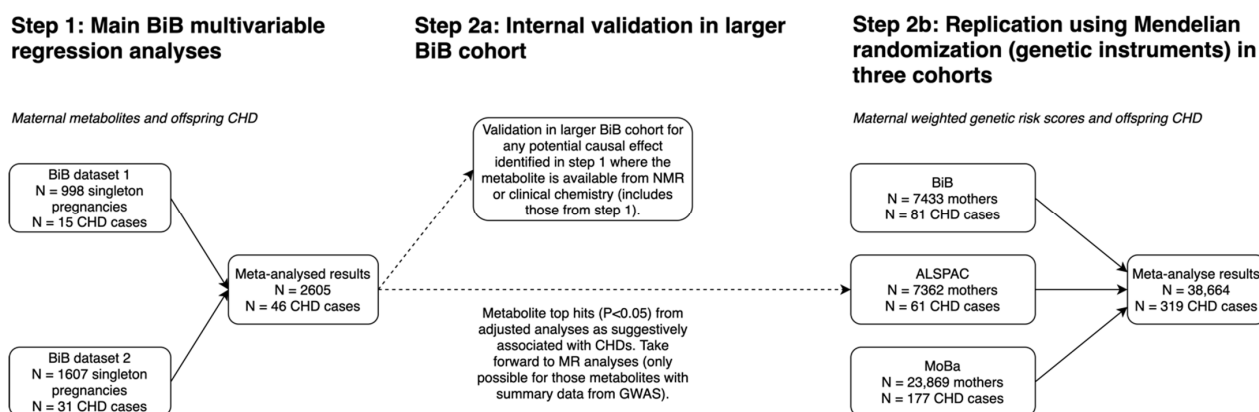


Figure 1. An overview of the study design. BiB has pregnancy mass spectrometry derived metabolomics in two separate datasets. Dataset 1 was completed in December 2017 and included 1000 maternal pregnancy samples. Dataset 2 was completed in December 2018 and consisted of 2000 maternal pregnancy samples within a case cohort design. The selection of participants into the two MS metabolomic datasets are shown in flowcharts in Figure S1. Abbreviations: CHD, congenital heart disease; BiB, Born in Bradford; NMR, nuclear magnetic resonance; MR, Mendelian randomization; GWAS, genome-wide association study; ALSPAC, Avon Longitudinal Study of Parents and Children; MoBa, Norwegian Mother, Father and Child Cohort.

2.2. Sample Collection and Metabolomic Profiling in BiB

Of the 13,776 pregnancies in the BiB cohort, 11,480 had a fasting blood sample taken during the OGTT (N = 10,574 [92%] between 26–28 weeks' gestation, with the remaining women being within 11–39 weeks' gestation). The samples were taken by trained phlebotomists working in the antenatal clinic of the Bradford Royal Infirmary and sent immediately to the hospital laboratory. The metabolomics data in the BiB cohort has previously been described in detail [15]. In brief, the metabolomics analysis was performed on ethylenediamine tetraacetic acid (EDTA) plasma samples at around 26–28 weeks' gestation. The untargeted MS metabolomics analysis of over 1000 metabolites was performed at Metabolon, Inc. (Durham, NC, USA). Quality control of the metabolite data was conducted by Metabolon. The classes of metabolites include amino acids, carbohydrates, cofactors and vitamins, energy, lipids, nucleotides, partially characterised molecules, peptides, and xenobiotics. These super-pathways, as defined by Metabolon, were also further subdivided into ~80 sub-pathways. The metabolite concentrations were quantified using the area under the curve of primary MS ions and were expressed as the multiple of the median (MoM) value for all batches processed on the given day. The MoM more closely reflects the biological variation rather than technical variation between samples or an analysis platform [25]. Due to the timing of funding acquisition, the samples were sent to Metabolon in two separate batches. Dataset 1 was completed in December 2017 and included 1000 maternal pregnancy samples. Dataset 2 was completed in December 2018 and consisted of 2000 maternal pregnancy samples within a case cohort design. The selection of the participants into the two MS metabolomic datasets are shown in flowcharts in Figure S1 and have been described in detail previously [15]. Neither dataset was random given how they were selected (Figure S1). The sampled cases of other adverse pregnancy/perinatal outcomes (that were not CHD) were removed.

2.3. Confounders

In multivariable regression analyses in BiB, we adjusted for the following maternal characteristics based on their known or plausible influence on maternal metabolites and on CHD: age, ethnicity, parity, residential neighbourhood index of multiple deprivation (IMD), body mass index (BMI), smoking, and alcohol consumption. Details of the methods for how the confounders were assessed are provided in the Supplementary File S1 (Text S1).

2.4. Congenital Heart Disease Outcomes

In the BiB cohort, the cases were identified from either the Yorkshire and Humber congenital anomaly register database, which tends to pick up most cases that were diagnosed antenatally and in the early postnatal period of life, or through linkage to primary care (up until age 5), which will have picked up any additional cases, in particular those that might have been less severe and not identified antenatally/in early life [26]. All the BiB cases were confirmed postnatally and were assigned ICD-10 codes. We used the ICD-10 codes to assign CHD cases according to the European surveillance of congenital anomalies (EUROCAT) guidelines. In the ALSPAC cohort, the cases were obtained from a range of data sources, including health record linkage and questionnaire data up until age 25 following European EUROCAT guidelines [27]. In the MoBa cohort, information on whether a child had a CHD or not (yes/no) was obtained through linkage to the Medical Birth Registry of Norway (MBRN). All maternity units in Norway must notify births to the MBRN, and information on malformations is reported to the registry up to 12 months postpartum [28]. Further details on defining CHDs including the ICD codes are shown in Text S2 (Supplementary File S1).

2.5. Genetic Data

The rationale for performing the MR analyses was to further explore potential causal mechanisms by using a method, namely, MR, which has different key sources of bias to multivariable regression. Metabolites are affected by multiple disease processes as well as numerous environmental exposures; therefore, understanding the metabolic pathways implicated in CHD is nontrivial. Genetic variants are less likely to be confounded by the socioeconomic and environmental factors that might bias causal estimates in conventional multivariable regression [29], but may be biased by weak instruments, which would bias towards the confounded association, or a path from the metabolomic genetic score to CHD, for example, via horizontal pleiotropy or foetal genotype [30]. Consistent results from both increase the confidence that the result is causal. In addition, whilst we included the BiB cohort in these genetic analyses, we also had two independent studies and repeated MR analyses with only the two (non-BiB) independent cohorts.

2.5.1. Genotyping in Each Cohort

The ALSPAC mothers were genotyped using an Illumina Human 660K-Quad single nucleotide polymorphism (SNP) chip, and the ALSPAC children were genotyped using the Illumina Human Hap 550 Quad genome-wide SNP genotyping platform. Genotype data for both the ALSPAC mothers and children were imputed against the Haplotype Reference Consortium v1.1 reference panel, after performing the QC procedure (with a minor allele frequency (MAF) $\geq 1\%$, a call rate $\geq 95\%$, in Hardy–Weinberg equilibrium (HWE), correct sex assignment, no evidence of cryptic relatedness, and of European descent). The samples of the BiB cohort (mothers and offspring) were processed on three different types of Illumina chips: HumanCoreExome12v1.0, HumanCoreExome12v1.1 and HumanCoreExome24v1.0. The genotype data were imputed against the UK10K + 1000 Genomes reference panel, after a similar QC procedure (with a call rate $\geq 99.5\%$, correct sex assignment, no evidence of cryptic relatedness, and correct ethnicity assignment). In MoBa, blood samples were obtained from both parents during pregnancy and from the mothers and children (umbilical cord) at birth [31]. Genotyping has had to rely on several projects—each contributing with resources to genotype subsets of MoBa over the last decade. The data used in the present study was derived from a cohort of genotypes samples from four MoBa batches. The MoBa genetics QC procedure involved a MAF $\geq 1\%$, a call rate $\geq 95\%$, was in HWE, included correct sex assignment, and had no evidence of cryptic relatedness. Further details of the genotyping methods for each cohort are provided in the Supplementary File S1 (Text S3) including flow charts showing the selection of the participants (Figure S2).

2.5.2. GWAS Data and SNP Selection

We aimed to construct weighted GRSs for all metabolites that were associated with CHD in the multivariable regression with a p -value < 0.05 (referred to throughout as “suggestively associated”) in the multivariable regression analyses using BiB data. To do this, we cross-referenced our suggestive associations with the relevant GWAS and we used summary data from two GWAS. In the first, the authors explored the genetic associations with 174 metabolites (compared with the 923 included in our study) [32]. To ensure robust associations, the SNPs were selected from the first GWAS if they were genome-wide significant ($p < 5 \times 10^{-8}$). Independent SNPs were defined as linkage disequilibrium (LD) $r^2 \geq 0.001$ assessed at a distance of 10,000 kb, with only one SNP from a correlated (clumped) group retained (performed using the TwoSampleMR package [33]). In the second (unpublished) GWAS, the authors performed a GWAS of 913 metabolite levels using samples from the EPIC-Norfolk [34] and INTERVAL studies [35]. A total of 14,296 participants were included in a discovery set (5841 from EPIC-Norfolk; 8455 from INTERVAL) and 5698 from EPIC-Norfolk in a validation set. The authors performed exact conditional analyses to identify independent associations.

2.5.3. Genetic Risk Score Generation

GRSs were calculated using SNPs previously associated in largescale GWAS with the metabolites (described above) by adding up the number of metabolite-increasing alleles among the selected SNPs after weighting each SNP by its effect on the corresponding metabolite:

$$GRS = w1 \times SNP1 + w2 \times SNP2 + \dots wn \times SNPn$$

where w is the weight (i.e., the beta-coefficient of the association of the SNP with the exposure from the published GWAS) and SNP is the genotype dosage of the exposure-increasing alleles at that locus (i.e., 0, 1, or 2 exposure-raising alleles). After matching the metabolites suggestively associated with CHDs at $p < 0.05$ from the multivariable regression analyses and removing indels, the selected SNPs were extracted from the imputed genotype data in dosage format using QCTOOL (v2.0) and VCF tools (v 0.1.12b) in ALSPAC and BiB, respectively. PLINK (v1.9) was then used to construct the GRS for each exposure coded so that an increased score was associated with increased levels in metabolites. In MoBa, we constructed the GRSs from the QC'd data in PLINK format. If a SNP was missing, a proxy SNP was used where available based on a $r^2 > 0.8$ using the European reference panel in the LDLink R package [36].

2.6. Statistical Analysis

Analyses were performed in R version 4.0.2 (R Foundation for Statistical Computing, Vienna, Austria). An analysis plan was written and uploaded to the Open Science Framework before the analyses commenced, where any subsequent changes to the analyses were documented along with the rationale [37]. We used scaled imputed data (in which missing data have been imputed and the multiple of median values transformed to standard deviation (SD) scores) which was log transformed. Any metabolite (in either dataset) where there was too little variation for meaningful analyses (defined as < 440 unique values) was excluded [38]. The transformed metabolite values were converted to standard deviation (SD) units. There were 1100 and 1150 quantified metabolites included in dataset 1 and 2, respectively, with 923 of these present in both datasets.

2.6.1. Multivariable Regression (Metabolomic) Analyses

We used logistic regression to estimate the odds ratios (ORs) and 95% confidence intervals (CIs) of any CHD per SD higher metabolite, with and without an adjustment for confounders. As we are interested in the potential causal effects, we present confounder adjusted results throughout. The analyses were completed separately in the two BiB

datasets and the results pooled using fixed-effects meta-analyses. Given that CHD is rare and binary, we accepted an uncorrected $p < 0.05$ (from meta-analyses) for the metabolite being suggestively associated with CHD in the offspring (but requiring further validation). We took these metabolites forward to the MR analyses.

The MS-platform used in BiB includes the measures of xenobiotics which are synthetic chemicals that are not synthesised by humans. Their presence in the circulation usually reflects endogenous exposures, such as medications and supplements. Given that these metabolites would not be present in all participants (and therefore have high missingness), many were removed (86/154 (56%)) from the dataset given the metabolite inclusion criteria of 440 unique observations mentioned above. Consequently, we performed an exploratory additional analysis using xenobiotics ($N = 154$) as binary variables (1 = yes, metabolite is detected in the sample; 0 = no, metabolite is not detected in the sample). We present adjusted ORs of these binary variables (any presence vs. none) with CHDs.

We sought to internally validate any of the metabolites suggestively associated with CHDs that were also measured in BiB in larger numbers using different methods. After matching the suggestive associations, we used data from the NMR platform ($N = 2$ metabolites) and did not use any data from the clinical chemistry measurements. More information on the BiB NMR data including the methods, QC and participant information has been described in detail previously [15].

2.6.2. Mendelian Randomization Analyses

We undertook MR in each of the 3 cohorts, including all BiB, ALSPAC, and MoBa participants with maternal genetic data and offspring CHD data. Logistic regression was used to estimate the OR of CHD per SD change in GRS, with an adjustment for the first 10 genetic principal components (PCs) with additional adjustments for the genetic chip, genetic batch, and imputation batch in MoBa.

The key assumptions for MR are: (i) Relevance assumption—the genetic instruments are robustly and strongly statistically associated with the exposure and relevant to the population being studied (i.e., here the pregnant women). We tested the association of the GRS of each metabolite with the metabolite levels during pregnancy in BiB dataset 2 using linear regression. (ii) Independence assumption—the IV outcome association is not confounded. Such confounding could occur as a result of population stratification. To minimise this, we adjusted the GRS-CHDs associations for the first 10 genetic PCs. We also repeated the MR analyses without the inclusion of BiB, given that BiB has a unique ethnic structure of South Asians and White Europeans. (iii) Exclusion restriction criteria—the genetic variant is not related to the outcome other than via its association with the exposure. We assessed pleiotropy by estimating the variance explained in all metabolites by each of the GRSs by undertaking the linear regression of every metabolite measured in BiB on each GRS. If the variance explained in other metabolites was similar or greater than to that explained in the candidate risk metabolite, this would suggest that there was low metabolite-specificity for the GRS and potential horizontal pleiotropic bias via the other metabolite(s). Importantly, however, this approach of testing the GRS specificity does not distinguish between vertical pleiotropy (e.g., the GRS influences the candidate metabolite which is the precursor of another metabolite that affects CHD) and horizontal pleiotropy (e.g., the GRS influences two metabolites that affect CHD independently). We also checked the consistency of the MR results when additionally adjusting for foetal genotype [30]. We performed the MR analyses separately in BiB, ALSPAC and MoBa and report the pooled results from random-effect meta-analyses for all three cohorts and fixed-effect meta-analyses for the MR analyses excluding BiB (i.e., ALSPAC and MoBa).

3. Results

3.1. Main BiB Multivariable Regression Analyses

Table 1 shows the distributions of characteristics for the women in both BiB datasets. In total, there were 2605 mother–offspring pairs with 46 CHD cases included in the BiB

multivariable regression metabolomic analyses. Table S1 shows the maternal characteristics stratified by offspring CHD status. There were some differences in the characteristics between the CHD and non-CHD cases in Dataset 1 such as a higher proportion of offspring of female sex and Pakistani ethnicity. The Dataset 2 characteristics were comparable between the CHD cases and non-CHD cases. N.B., for consistency and clarity, we refer to metabolites here by their super-pathways (as defined by Metabolon). A metabolite might have a different super-pathway and chemical group. For example, N-Acetylcarnosine is a metabolite that is part of the amino acid super-pathway, but it is not an amino acid itself. Where available, we include the Human Metabolome Database (HMDB) IDs with all numerical results, which can assist the reader in finding further information on the structure and function of a metabolite. The super-pathways that included the largest proportions of the 923 metabolites were the lipids (38%), unknown (22%), amino acids (18%) and Xenobiotics (8%), with other super-paths having $\leq 3\%$ of the total (Table 2).

Table 1. Participant characteristics for the Born in Bradford metabolomic analyses.

Characteristic	Category	BiB Dataset 1 (N = 998)	BiB Dataset 2 (N = 1607)
<i>Offspring</i>			
CHD	Yes	15 (1.6)	31 (1.9)
Sex	Male	510 (51.1)	844 (52.5)
	Female	488 (48.9)	763 (47.5)
<i>Maternal</i>			
Age, years		27.5 (5.7)	27.3 (5.6)
Parity	Nulliparous	358 (37.0)	616 (36.8)
	Multiparous	610 (63.0)	991 (63.2)
BMI, kg/m ²		26.7 (6.0)	26.5 (5.8)
Ethnicity	White British	500 (50.0)	733 (45.6)
	Pakistani	498 (50.0)	874 (54.4)
Neighbourhood deprivation (IMD)	Quintile 1 (most deprived)	654 (65.5)	1084 (67.5)
	Quintile 2	175 (17.5)	281 (17.5)
	Quintile 3	112 (11.2)	175 (10.9)
	Quintile 4	38 (3.8)	40 (2.5)
	Quintile 5 (least deprived)	19 (1.9)	27 (2.7)
Smoking	Yes	176 (17.7)	311 (19.2)
Alcohol	Yes	338 (33.9)	496 (30.8)
Gest age at blood sampling, weeks		26.2 (2.0)	26.2 (2.0)

Data are means \pm SD or n (%) unless stated. Abbreviations: BiB, Born in Bradford; CHD, congenital heart disease; BMI, body mass index; kg, kilogram; m, meter; HDP, hypertensive disorders of pregnancy; GHT, gestational hypertension; PE, pre-eclampsia; IMD, Index of Multiple Deprivation (taken from 2010 national quintiles); gest, gestational.

Table 2. Showing the breakdown of metabolites in our dataset (N = 923) into the 10 super-pathways as defined by Metabolon.

Super Pathway	N (%) for All Metabolites (N = 923)	N (%) for Metabolites Suggestively Associated with CHDs (N = 44)
Amino Acid	170 (18.4%)	10 (22.7%)
Lipid	354 (38.4%)	18 (40.9%)
Cofactors and Vitamins	27 (2.9%)	4 (9.0%)
Partially Characterized Molecules	3 (0.3%)	1 (2.3%)
Unknown	201 (21.8%)	7 (15.9%)
Xenobiotics	86 (9.3%)	2 (4.5%)
Nucleotide	33 (3.6%)	1 (2.3%)
Energy	8 (0.9%)	1 (2.3%)
Carbohydrate	19 (2.1%)	0
Peptide	22 (2.4%)	0

Abbreviations: CHD, congenital heart disease.

Of the 923 metabolites quantified in both BiB datasets, 44 (4.8%) were associated with any CHD, at $p < 0.05$, in the confounder adjusted pooled analyses (Figure 2). We observed suggestive effects (i.e., confounder adjusted associations reaching the p -value threshold < 0.05) with several amino acids, lipids and co-factors and vitamins. There were also suggestive effects for two xenobiotics, one nucleotide, one energy metabolite and some partially characterised and unknown metabolites (Figure 2). None of the 22 peptide or 19 carbohydrate-related metabolites associated with CHD at this p -value threshold. Of the 18 lipid-related metabolites associated with CHD, 13 were positively associated (i.e., increased odds) (e.g., Glycolithocholate Sulfate: adjusted odds ratio (aOR) per SD increase in metabolite: 1.73 95% CI (1.21, 2.48)) and 5 were negatively associated (decreased odds) (e.g., Phosphocholine: aOR 0.65 (0.47, 0.90)). All but one (N-Acetylcarnosine) of the 10 amino acid-related metabolites were negatively associated with CHDs (e.g., isoleucine: aOR: 0.67 (0.49, 0.92)). Three of the four co-factors and vitamins were negatively associated, whereas one (biliverdin) was positively associated (aOR 1.41 (1.07, 1.86)). One nucleotide was negatively associated (inosine 5'-Monophosphate (Imp): aOR 0.59 (0.36, 0.99)) and one energy related metabolite was positively associated (succinylcarnitine (C4): aOR 1.42 (1.02, 1.97)). Benzoate and Saccharin were the two xenobiotics associated with CHDs in the main analyses with both showing positive associations. The results for the associations of all metabolites (irrespective of p -value) in the unadjusted and confounder adjusted analyses from the pooled datasets, and each dataset separately are provided in Supplementary File S2 (Tables S5–S7), including HMDB IDs where applicable.

In the analysis treating the xenobiotics as binary variables, after removal of the metabolites with no exposed cases, there were six xenobiotic metabolites suggestively associated with offspring CHDs (Supplementary File S1: Table S2). Two out of the six showed positive associations: saccharin, which was also associated in the main analyses (adjusted odds ratio (aOR) for the presence of metabolite vs. not: 2.16 95% CI (1.02, 5.13)—an artificial sweetener) and salicyluric glucuronide (aOR: 2.27 (1.16, 4.29)—a metabolite involved in aspirin metabolism). The remaining four showing negative associations are all part of the food component/plant metabolite sub-pathway (Table S2).

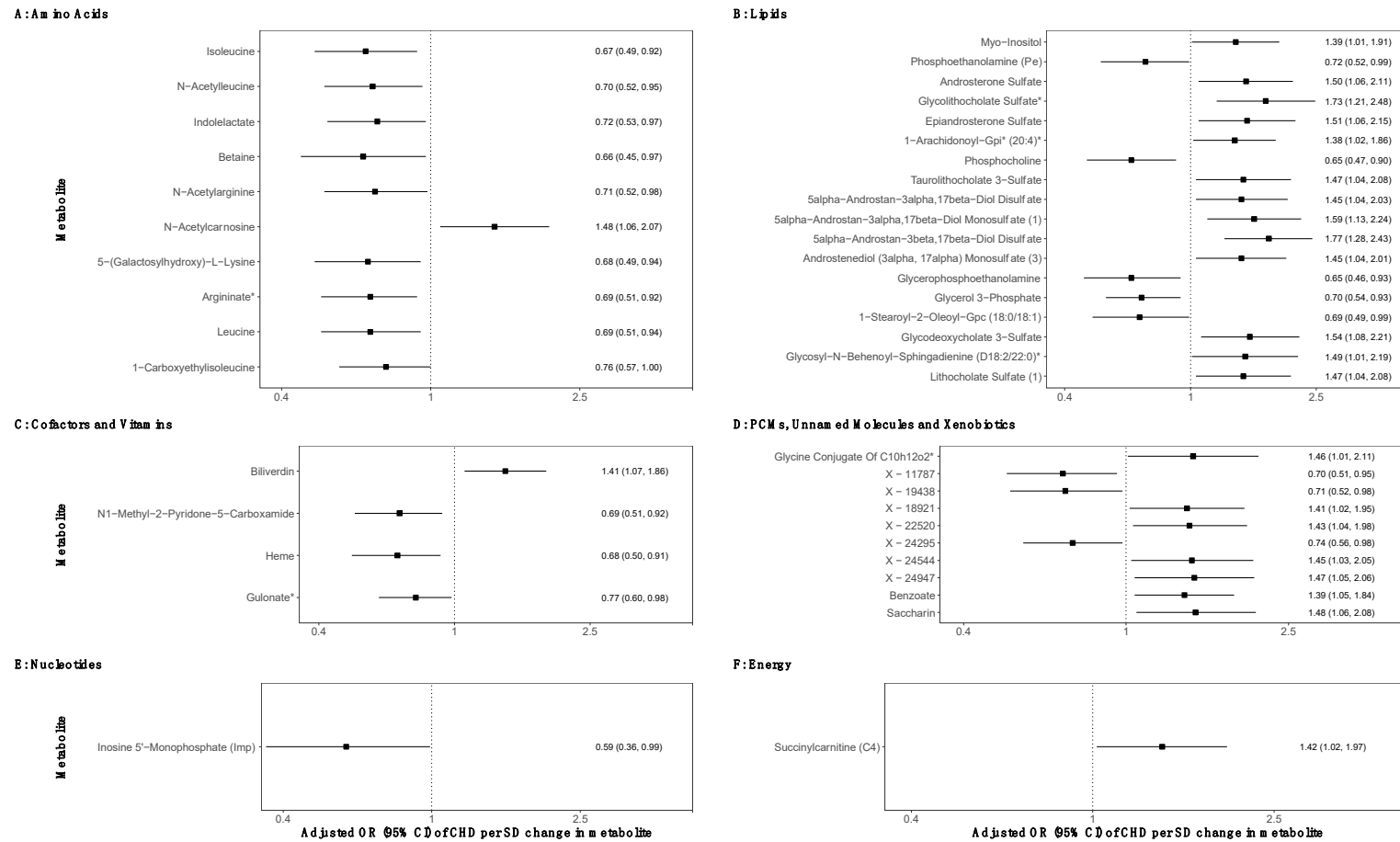


Figure 2. Pooled confounder adjusted associations of maternal pregnancy metabolites with offspring congenital heart disease in the Born in Bradford cohort (N = 2391 and N CHD cases = 42). The associations show confounder adjusted odds ratios of CHD per standard deviation change in log-transformed metabolite levels for the 44 (out of 923) metabolites that associated with CHD at p -value < 0.05 separated by super pathways as defined by Metabolon. Metabolites were measured at ~26–28 weeks’ gestation. Heterogeneity statistics and separate associations for datasets 1 and 2 are reported in Supplementary Tables S5–S7. Associations were adjusted for maternal age, ethnicity, parity, Index of Multiple Deprivation, body mass index, smoking and alcohol intake. * at the end of metabolite names as defined by Metabolon. Abbreviations: PCMs, partially characterised molecules; OR, odds ratio; CHD, congenital heart disease; SD, standard deviation.

3.2. Internal Validation Using NMR or Clinical Chemistry Measures of Suggestive Associations from Main Multivariable Regression Analyses

It was possible to explore 2 of the 44 metabolites suggestively associated with CHDs in the larger BiB sample. In the comparable confounder adjusted analyses, the NMR measured amino acids isoleucine and leucine were available on 7296 mothers, with 87 having an offspring with CHD. The results for these two amino acids were highly consistent between the two samples/assay methods (aOR per SD increase in the MS isoleucine 0.67 (0.49, 0.92) vs. 0.65 (0.50, 0.84) for the NMR isoleucine, and aOR per SD increase in the MS leucine 0.69 (0.51, 0.94) vs. 0.67 (0.53, 0.85) for NMR leucine).

3.3. Validating Findings with Mendelian Randomization

The distributions of offspring and maternal characteristics for the MR analyses in BiB, ALSPAC and MoBa are displayed in Table S3 (Supplementary File S1). It was possible to explore MR replication for 27 of the 44 metabolites that associated with CHD in the multivariable analyses (the other 17 were either not available in the GWAS or had genetic variants not available in the cohorts, Figure S3). All but three of the GRSs (24/27 (89%)) were associated with the corresponding metabolite during pregnancy in BiB (with R^2 values ranging from 0.3% to 34%), but for the remaining three, the associations were imprecise with wide confidence intervals that included the null (Table S4; Supplementary File S1). Of the 27 GRSs, 3 were specific for the metabolite they were instrumenting (i.e., they had the strongest association with it and little evidence of associations with other metabolites; N-acetylcarnosine, phosphocholine and succinylcarnitine). Eighteen GRSs were associated with the metabolite they were instrumenting and several others that were correlated with that metabolite (e.g., the biliverdin GRS was associated with it and also similarly with other hepatic-related metabolites). Six GRSs were more strongly associated with other (uncorrelated) metabolites than the one they were instrumenting (scatter plots for all twenty-seven GRSs are shown in Figure S4; Supplementary File S1). The six non-specific GRS were for indolelactate, glycolithocholate sulfate, isoleucine, leucine, myo-inositol and tauroolithocholate 3-sulfate (the MR results for these should be treated with caution and are denoted in Figure 3B by white-filled points).

Results from the MR analyses demonstrated the same direction of (protective) effect of higher levels of the amino acids leucine, indolelactate and isoleucine on CHD (although the instruments for these metabolites were non-specific as noted above), but for the other amino acids, the MR results were either very close to the null or in the opposite direction (Figure 3). Seven of the lipid-related metabolites that were positively associated in multivariable regression were also replicated in MR analyses (six of which were highly correlated androgenic steroid metabolites), as was the energy related metabolite, succinylcarnitine (Figure 3). For the 11 metabolites where we considered the MR GRS analyses providing some evidence of replication and a potential causal effect, 7 of the GRSs were specific for the metabolite alone and/or also for its correlates. Individual study results and the p -values for heterogeneity are included in Supplementary File S2 (Table S8). The MR results were largely unchanged when excluding BiB from the analyses (Table S8) and when adjusting for the offspring genotype (Table S9; Supplementary File S2).

Exploring directional consistency between phenotype (conventional multivariable regression) and genotype (Mendelian randomization) associations with metabolites

A: Main multivariable regression results

B: Mendelian randomization results

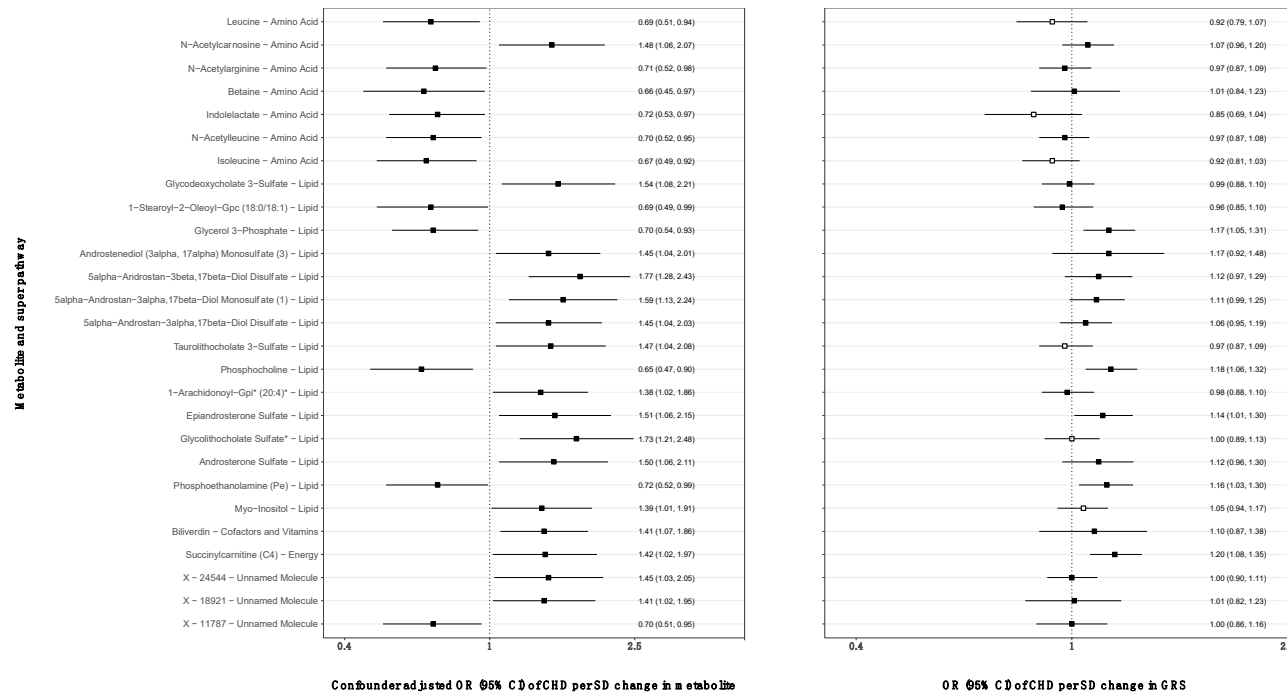


Figure 3. Exploring directional consistency between phenotype (conventional multivariable regression) and genotype (Mendelian randomization) associations with metabolites. Showing results comparing the main confounder adjusted associations of maternal metabolites with offspring CHDs (Panel A: N = 2391 and N CHD cases = 42 in the Born in Bradford cohort) to the Mendelian randomization analyses of maternal genetic risk scores and offspring CHDs (Panel B: N = 38,662 and N CHD cases = 319 across 3 cohorts). N.B., results from each analysis are presented on different scales; we are not attempting to quantify estimates in the MR analyses, the aim is to compare the direction of effect. The confounder adjusted associations are as above in Figure 2. The MR analyses are adjusted for the top 10 genetic principal components and genetic batches in MoBa. In Panel B, the metabolite genetic risk scores filled with white appeared to be non-specific for the metabolite we were trying to instrument (i.e., the risk score relates to several other metabolites more strongly than the specific named metabolite). The metabolites filled in black were either metabolite-specific or specific to the metabolite and other correlated metabolites (see scatter plots in Figure S4). The results were pooled using random effects meta-analyses; individual study results and p-values for heterogeneity are shown in Supplementary Table S8. * at the end of metabolite names as defined by Metabolon. Abbreviations: BiB, Born in Bradford; CHD, congenital heart disease; GRS, genetic risk score; MR, Mendelian randomization; OR, odds ratio; CI, confidence interval.

4. Discussion

Maternal metabolism is important for healthy foetal growth and development. To our knowledge, no previous study has examined the association of detailed maternal metabolites with the risk of CHD within a causal framework. In this study, we found 44 metabolites (of 923) suggestively associated with CHD. These included metabolites related to amino acids, lipids, co-factors and vitamins, unknown molecules, xenobiotics, nucleotides and energy. In separate xenobiotics analyses, there was some evidence that metabolites related to aspirin and saccharine may increase the odds of CHD, whereas metabolites related to plant food components may reduce the odds. Two of the amino acids, where it was possible to explore replication, were replicated in BiB in larger numbers using an alternative metabolomics platform. In the MR analyses, we analysed the relationship of maternal genetic variants that are associated with metabolite levels with offspring CHD. We completed this for 27 metabolites that showed evidence of association with offspring CHD in multivariable regression analyses and that we had GWAS summary data on. There was directional consistency for 11 out of the 27 metabolites that could be explored in the MR analyses. Overall, we found that maternal amino acid metabolism during pregnancy, several lipids (more specifically androgenic steroids), and levels of succinylcarnitine could be important contributing factors to offspring CHD risk.

Nine out of the ten amino acids suggestively associated with CHDs were negatively associated, suggesting that deficiencies in certain amino acids during pregnancy could contribute to offspring CHDs. Previous research found that amino acid concentrations measured in amniotic fluid were lower in patients with CHDs [39], which is a similar pattern to what we found here. We were able to replicate our findings for isoleucine and leucine in larger numbers in BiB, which improves the confidence in these findings. The MR analyses also provided evidence to support the direction of association for these metabolites; however, the GRSs for isoleucine and leucine were non-specific, and, therefore, these results should be treated with caution until these genetic instruments have been shown to be robustly associated with the amino acid metabolites.

Eighteen of the forty-four maternal metabolites suggestively associated with CHD were part of the lipid super pathway. Previous work reported that an abnormal lipid profile (defined as elevated cholesterol and apolipoprotein B) [5], abnormal lipid metabolism (defined as a disturbance in phosphatidyl-choline and various sphingolipids and choline metabolism) [13] and high maternal blood lipids [6] are a feature of CHD pregnancies. We were able to take forward 15 (out of 18) of the lipid metabolites and replicated the direction of effect for 7. All except one of these seven replicated metabolites were androgenic steroids and, therefore, were highly correlated. Steroids are important for numerous functions during gestation, particularly for normal placental function [40]. Here we present evidence of a potential causal effect (associated in metabolomic analyses with consistent direction of effect in the MR analyses) of positive associations between maternal gestational androgenic steroid metabolites and offspring CHDs.

Levels of bilirubin and biliverdin were positively associated with CHDs. MR analyses were only possible for biliverdin but were inconclusive with wide confidence intervals; thus, these two compounds, which are involved in heme catabolism, should be further investigated for a possible role in CHD development. Levels of succinylcarnitine were also positively associated with CHDs and we found good replications in the MR analyses with consistent directions of effect and a GRS that appeared highly specific for succinylcarnitine. Succinylcarnitine is an acylcarnitine, which is a group of metabolites responsible for the beta oxidation of fatty acids and mitochondrial function [41]. It is well documented that fatty acids play an important role in embryonic and foetal development [42,43]. We included the analyses of partially characterised and unknown metabolites in our results as with the increasing evidence from genomic studies, previously unknown metabolites are having their function identified. With future studies identifying the function of some of these unknown/partially characterized metabolites, our results could shed light on the aetiology of CHDs.

A key strength of this study is the unique data that was available in BiB to support novel analyses of the associations of a wide range of maternal metabolic paths with offspring CHD risk. We were not able to identify any other study with such data; however, we realised, even before the analyses, that we would have limited statistical power with just 46 CHD cases. This motivated us to think about ways of trying to replicate any findings in larger samples, either through finding measures of the same metabolites available from other assays in larger samples or using GRSs as instruments for the metabolites. In the initial multivariable regression analyses we adjusted for potential confounders. We defined the suggestive associations based on a p -value threshold < 0.05 , i.e., not taking account of multiple testing. When we applied a Bonferroni corrected threshold ($p < 0.0001$), none of the associations passed this (Supplementary Table S7; Supplementary File S2). Given the novel nature of this study and the use of the initial multivariable regression in BiB to select associations for further follow-up (replication and MR), we felt this was appropriate. As with any 'screening' for further analyses, we wanted to ensure that we would not miss potential causal effects. We recognise that selecting the results based on a p -value threshold is problematic, as some associations with higher p -values might have associations of a magnitude that could be clinically important, but there would also be the potential for several false positives. Additionally, we limited the MR analyses only to those metabolites that associated with a $p < 0.05$ rather than undertaking these analyses on all of the 923 metabolites. Our reason for this was that having searched for all studies with maternal genome wide data and offspring CHD outcomes, we identified only three cohorts and recognised that for MR analyses, the pooled results from these might also have limited power. The limited power in both the multivariable and MR analyses also meant that we could only examine associations with any CHD and not subtypes.

The MR analyses were sensitive to their assumptions that the GRS is statistically strongly associated with the metabolite in pregnancy. We examined the associations of these with pregnancy metabolite levels and were careful in our interpretation of the results in relation to this. Methods that are available for exploring potential bias due to horizontal pleiotropy in two-sample MR were not possible here. We know that many of the 923 metabolites will be biologically related to each other and with our sample size it would be difficult to robustly distinguish the effects of correlated metabolites. We explored this by examining the strength of association (proportion of the variation explained) of each of the 27 GRSs with all other metabolites available in the BiB dataset 2. Stronger or similar associations with other metabolites would suggest that the GRS was not a specific instrument for the metabolite that we were using it for. In this case this could be because of known biological relations. For example, we know biologically that many of the lipid metabolites are related to each other, and we saw this with similar proportions of variation explained by the GRS of the androgenic steroid lipid metabolites with other androgenic steroid lipid metabolites. As such, we would interpret the results for these metabolites as supporting an effect of maternal androgenic steroid metabolites on CHD, but we could not be specific about which ones were driving this. A similar or stronger variation of a GRS for other metabolites could be related to vertical pleiotropy, i.e., the metabolite for which the GRS is instrumenting strongly influences other metabolites that are related to CHD with the other metabolites partly mediating the effect of the focused metabolite. This would not bias the result; however, this could also occur with horizontal pleiotropy where the GRS, independently of the metabolite of interest, influences other metabolites that are risk factors for CHD. With our current data, we are not able to distinguish between these two.

A further limitation of this study is that the maternal plasma/serum metabolomics data were derived at a single timepoint around 26–28 weeks' gestation. Foetal cardiac development starts early in pregnancy and much of the development occurs in the first trimester [44]. Here, we are assuming that metabolite levels around 26–28 weeks' gestation are good proxies for levels in early pregnancy—when the offspring's heart is forming. Previous work has shown that between person differences throughout pregnancy remain largely consistent (i.e., those with a high level of a metabolite in early pregnancy tend to

have a similarly high level of a metabolite in later pregnancy) [45]. Similarly, and worth mentioning, the effects obtained from MR studies are often interpreted as the lifetime effect of the exposure (metabolites) in question [46]. We were also not able to adjust for diet in the multivariable regression analyses. Metabolite concentrations are likely to be influenced by differences in dietary intake and there is some evidence that differences in maternal pregnancy diets may influence the risk of CHD [47]; thus, our multivariable regression analyses may be influenced by residual confounding due to diet or other factors. The selection process to include women in the BiB metabolomic multivariable regression analyses may have also influenced (in either direction) the results through selection bias. Again, where it was possible to explore replication in the larger BiB sample with the metabolites measured using NMR or clinical chemistry and/or undertaking MR, the inclusion of more of the BiB cohort and the use of two independent cohorts were useful, as these results are less likely to be influenced by selection.

In summary, we have used metabolomics and genetic data obtained during pregnancy to explore how the maternal metabolome may contribute to offspring CHDs. We found evidence that markers of maternal pregnancy amino acids and androgenic steroids during pregnancy, as well as higher levels of succinylcarnitine could be important contributing factors to offspring CHD risk. It was not possible to undertake genetic MR analyses on 17 of the potential associations from multivariable regression and further multivariable regression and/or MR in larger independent samples, but these would be valuable. Our analysis pipeline, which involved seeking replication of metabolite associations by harnessing large-scale GWAS data, provides scope to address this and to obtain more precise estimates for the 27 associations we could undertake MR on, as more cohorts with genomic and metabolomic data become available. Metabolomics could prove to be an important tool for identifying the biological pathways that may lead to the identification of prevention targets to decrease the disease burden of CHDs. To do this, future research will require the international collaboration of more and larger studies with detailed metabolomics data in pregnancy, ideally with some of these having repeat measures across pregnancy and offspring CHD data.

4.1. Ethical Approval and Consent to Participate

Ethical approval for ALSPAC was obtained from the ALSPAC Law and Ethics committee and local research ethics committees (NHS Haydock REC: 10/H1010/70). Informed consent for the use of data collected via questionnaires and clinics was obtained from the participants following the recommendations of the ALSPAC Ethics and Law Committee at the time. At age 18, the study's children were sent 'fair processing' materials describing ALSPAC's intended use of their health and administrative records and were given clear means to consent or object via a written form. Data were not extracted for the participants who objected, or who were not sent fair processing materials. For BiB, Ethics approval has been obtained for the main platform study and all of the individual sub studies from the Bradford Research Ethics Committee. Written consent was obtained from all participants. The establishment of MoBa and the initial data collection was based on a license from the Norwegian Data Protection Agency and approval from The Regional Committees for Medical and Health Research Ethics. The MoBa cohort is now based on regulations related to the Norwegian Health Registry Act.

4.2. Availability of Data and Materials

The ALSPAC data management plan (<http://www.bristol.ac.uk/alspac/researchers/data-access/documents/alspac-data-management-plan.pdf>), describes in detail the policy on data sharing, which is through a system of managed open access. Scientists are encouraged to make use of the BiB study data, which are available through a system of managed open access. Please note that the study website contains details of all the data that is available through a fully searchable data dictionary and variable search tool and reference the following webpage: <http://www.bristol.ac.uk/alspac/researchers/our-data/>.

Before you contact the BiB study, please make sure you have read the Guidance for Collaborators: <https://borninbradford.nhs.uk/research/guidance-for-collaborators/>. MoBa data are used by researchers and research groups at both the Norwegian Institute of Public Health and other research institutions nationally and internationally. The research must adhere to the aims of MoBa and the participants' given consent. All use of data and biological material from MoBa is subject to Norwegian legislation. More information can be found on the study website (<https://www.fhi.no/en/studies/moba/for-forskere-artikler/research-and-data-access/>).

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/jcdd9080237/s1>, **Supplementary File S1:** Figure S1. Illustrating the flow of participants into the Metabolon datasets in the Born in Bradford cohort, Text S1. Confounder data, Text S2. Defining congenital heart disease, Text S3. Genetic data methods, Figure S2. An overview of included cohorts and selection of study participants for Mendelian randomization analyses, Table S1: Participant characteristics for the Born in Bradford metabolomic analyses stratified by CHD status, Table S2: Logistic regression results from the presence/absence xenobiotic analysis, Table S3. Participant characteristics for the 3 studies included in Mendelian randomization analyses, Figure S3. Flow chart to illustrate the analysis pipeline and selection of metabolites for Mendelian randomization, Table S4. Characteristics of maternal GRS and associations with the corresponding metabolite in BiB Dataset 2 (N = 1326), Figure S4. Scatter plots of the variance explained by the weighted genetic risk score for each of the 27 metabolites included in MR analyses. **Supplementary File S2:** Table S5. Associations of maternal pregnancy metabolites measured at around 26–28 weeks' gestation with offspring congenital heart disease (Dataset 1), Table S6. Associations of maternal pregnancy metabolites measured at around 26–28 weeks' gestation with offspring congenital heart disease (Dataset 2), Table S7. Associations of maternal pregnancy metabolites measured at around 26–28 weeks' gestation with offspring congenital heart disease, Table S8. Mendelian randomization results for each study and pooled across studies, Table S9. Mendelian randomization results pooled across studies to check consistency of results when adjusting for offspring genotype.

Author Contributions: Conceptualization, K.T. and D.A.L.; methodology, K.T. and D.A.L.; software, K.T.; formal analysis, K.T.; investigation, K.T.; resources, J.Z., S.O, R.A., J.W., O.A.A., I.D.S., C.L., M.C.B., M.C.M., M.C. and D.A.L.; data curation, K.T., I.D.S. and M.C.M.; writing—original draft preparation, K.T., N.M. and D.A.L.; writing—review and editing, All Authors; visualization, K.T.; supervision, M.C. and D.A.L.; project administration, K.T.; funding acquisition, J.W., O.A.A., C.L., M.C.M., M.C. and D.A.L. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the European Research Council under the European Union's Seventh Framework Programme (FP/2007–2013)/(grant agreement No 669545), US National Institute of Health (R01 DK10324) British Heart Foundation (CS/16/4/32482 and AA/18/7/34219) Taylor is supported by a British Heart Foundation Doctoral Training Program (FS/17/60/33474). Taylor, McBride, Borges, and Lawlor work in a unit that is supported by the University of Bristol and UK Medical Research Council (MC_UU_00011/6). Lawlor's contribution is also supported by a British Heart Foundation Chair in Cardiovascular Science and Clinical Epidemiology (CH/F/20/90003) and a NIHR Senior Investigator (NF-0616-10102). Caputo is supported by the British Heart Foundation Chair in Congenital Heart Disease (CH/1/32804). Magnus has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 947684). This research was also supported by the Research Council of Norway through its Centres of Excellence funding scheme (Project No. 262700). M.C. Borges is supported by a Vice Chancellor's fellowship from the University of Bristol. The BiB study has received core support from Wellcome Trust (223601/Z/21/Z and WT101597MA), a joint grant from the UK Medical Research Council and UK Economic and Social Science Research Council (MR/N024397/1), the British Heart Foundation (CS/16/4/32482), and the NIHR Applied Research Collaboration Yorkshire and Humber (NIHR200166) and Clinical Research Network. Core funding for ALSPAC is provided by the UK Medical Research Council and Wellcome (217065/Z/19/) and the University of Bristol. Many grants have supported different data collections, including for some of the data used in this publication, and a comprehensive list of grant funding is available on the ALSPAC website (<http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>) GWAS data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and

LabCorp (Laboratory Corporation of America) using support from 23andMe. The views expressed in this publication are those of the author(s) and not necessarily those of the National Institute for Health Research or the Department of Health and Social Care. The Norwegian Mother, Father and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. The EPIC-Norfolk study (<https://doi.org/10.22025/2019.10.105.00004>) has received funding from the Medical Research Council (MR/N003284/1 MC-UU_12015/1 and MC_UU_00006/1) and Cancer Research UK (C864/A14136). The genetics work in the EPIC-Norfolk study was funded by the Medical Research Council (MC_PC_13048). Metabolite measurements in the EPIC-Norfolk study were supported by the MRC Cambridge Initiative in Metabolic Science (MR/L00002/1) and the Innovative Medicines Initiative Joint Undertaking under EMIF grant agreement no. 115372.

Institutional Review Board Statement: See Section 4.1 above (ethical approval and consent to participate).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: See Section 4.2 above (availability of data and materials).

Acknowledgments: The BiB (Born in Bradford) study is only possible because of the enthusiasm and commitment of the children and parents. We are grateful to all the participants, practitioners, and researchers who have made the BiB study happen. We are extremely grateful to all the families who took part in The Avon Longitudinal Study of Parents and Children (ALSPAC) study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The authors thank all participating families in Norway who take part in this ongoing cohort study, and those who contributed to the recruitment and the infrastructure of the cohort. We also thank the Norwegian Institute of Public Health for generating high-quality genomic data and the NORMENT Centre for providing genotype data, South East Norway Health Authority, and Stiftelsen Kristian Gerhard Jebsen. The authors further thank the Center for Diabetes Research (University of Bergen) for providing genotype information and performing quality control and imputation of the data in research projects funded by the European Research Council Advanced Grant SELECTionPREDISPOSED, Stiftelsen Kristian Gerhard Jebsen, the Trond Mohn Foundation, the Research Council of Norway, the Novo Nordisk Foundation, the University of Bergen, and the Western Norway Health Authority. This work was performed on the Tjeneste for Sensitive Data (TSD) facilities, owned by the University of Oslo, operated and developed by the TSD service group at the University of Oslo, IT-Department (tsd-drift@usit.uio.no). This study does not necessarily reflect the position or policy of the Norwegian Research Council. We are grateful to all the participants who have been part of the EPIC-Norfolk metabolomics project and to the many members of the study teams at the University of Cambridge who have enabled this research.

Conflicts of Interest: D.A.L. has received support from Medtronic Ltd. and Roche Diagnostics for biomarker research unrelated to those presented in this paper. O.A.A. is a consultant to HealthLytix. All other authors declare they have not conflict of interest.

References

1. Wang, H.; Naghavi, M.; Allen, C.; Barber, R.M.; Bhutta, Z.A.; Carter, A.; Casey, D.C.; Charlson, F.J.; Chen, A.Z.; Coates, M.M.; et al. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015. *Lancet* **2016**, *388*, 459–544. Available online: <https://linkinghub.elsevier.com/retrieve/pii/S0140673616310121> (accessed on 20 February 2020). [CrossRef]
2. Blue, G.M.; Kirk, E.P.; Sholler, G.F.; Harvey, R.P.; Winlaw, D.S. Congenital heart disease: Current knowledge about causes and inheritance. *Med. J. Aust.* **2012**, *197*, 155–159. Available online: <https://onlinelibrary.wiley.com/doi/abs/10.5694/mja12.10811> (accessed on 5 February 2020). [CrossRef]
3. Monni, G.; Atzori, L.; Corda, V.; Dessolis, F.; Iuculano, A.; Hurt, K.J.; Murgia, F. Metabolomics in Prenatal Medicine: A Review. *Front. Med.* **2021**, *8*, 645118. Available online: <https://www.frontiersin.org/articles/10.3389/fmed.2021.645118/full> (accessed on 26 July 2021). [CrossRef]
4. Koster, M.P.H.; van Duijn, L.; Krul-Poel, Y.H.M.; Laven, J.S.; Helbing, W.A.; Simsek, S.; Steegers-Theunissen, R.P.M. A compromised maternal vitamin D status is associated with congenital heart defects in offspring. *Early Hum. Dev.* **2018**, *117*, 50–56. [CrossRef]
5. Smedts, H.P.M.; van Uitert, E.M.; Valkenburg, O.; Laven, J.S.E.; Eijkemans, M.J.C.; Lindemans, J.; Steegers, E.A.P.; Steegers-Theunissen, R.P.M. A derangement of the maternal lipid profile is associated with an elevated risk of congenital heart disease in

- the offspring. *Nutr. Metab. Cardiovasc. Dis.* **2012**, *22*, 477–485. Available online: <https://linkinghub.elsevier.com/retrieve/pii/S0939475310001900> (accessed on 27 March 2020). [CrossRef]
6. Cao, L.; Du, Y.; Zhang, M.; Wang, F.; Zhao, J.Y.; Ren, Y.Y.; Gui, Y.H. High maternal blood lipid levels during early pregnancy are associated with increased risk of congenital heart disease in offspring. *Acta Obstet. Gynecol. Scand.* **2021**, *100*, 1806–1813. [CrossRef]
 7. Priest, J.R.; Yang, W.; Reaven, G.; Knowles, J.W.; Shaw, G.M. Maternal Midpregnancy Glucose Levels and Risk of Congenital Heart Disease in Offspring. *JAMA Pediatr.* **2015**, *169*, 1112–1116. [CrossRef]
 8. Øyen, N.; Diaz, L.J.; Leirgul, E.; Boyd, H.A.; Priest, J.; Mathiesen, E.R.; Quertermous, T.; Wohlfahrt, J.; Melbye, M. Prepregnancy Diabetes and Offspring Risk of Congenital Heart Disease: A Nationwide Cohort Study. *Circulation* **2016**, *133*, 2243–2253. [CrossRef]
 9. Simeone, R.M.; Devine, O.J.; Marcinkevage, J.A.; Gilboa, S.M.; Razzaghi, H.; Bardenheier, B.H.; Sharma, A.J.; Honein, M.A. Diabetes and congenital heart defects: A systematic review, meta-analysis, and modeling project. *Am. J. Prev. Med.* **2015**, *48*, 195–204. [CrossRef]
 10. Bahado-Singh, R.O.; Ertl, R.; Mandal, R.; Bjorndahl, T.C.; Syngelaki, A.; Han, B.; Dong, E.; Liu, P.B.; Alpay-Savasan, Z.; Wishart, D.S.; et al. Metabolomic prediction of fetal congenital heart defect in the first trimester. *Obstet. Gynecol. Surv.* **2015**, *70*, 9–11. [CrossRef]
 11. Xie, D.; Luo, Y.; Xiong, X.; Lou, M.; Liu, Z.; Wang, A.; Xiong, L.; Kong, F.; Wang, Y.; Wang, H. Study on the Potential Biomarkers of Maternal Urine Metabolomics for Fetus with Congenital Heart Diseases Based on Modified Gas Chromatograph-Mass Spectrometer. *Biomed. Res. Int.* **2019**, *2019*, 1905416. [CrossRef]
 12. Li, Y.; Sun, Y.; Yang, L.; Huang, M.; Zhang, X.; Wang, X.; Guan, X.; Yang, P.; Wang, Y.; Meng, L.; et al. Analysis of Biomarkers for Congenital Heart Disease Based on Maternal Amniotic Fluid Metabolomics. *Front. Cardiovasc. Med.* **2021**, *8*, 671191. [CrossRef]
 13. Jaddoe, V.W.V.; Felix, J.F.; Andersen, A.M.N.; Charles, M.A.; Chatzi, L.; Corpeleijn, E.; Donner, N.; Elhakeem, A.; Eriksson, J.G.; Foong, R.; et al. The LifeCycle Project-EU Child Cohort Network: A federated analysis infrastructure and harmonized data of more than 250,000 children and parents. *Eur. J. Epidemiol.* **2020**, *35*, 709–724. [CrossRef]
 14. Wright, J.; Small, N.; Raynor, P.; Tuffnell, D.; Bhopal, R.; Cameron, N.; Fairley, L.; Lawlor, D.A.; Parslow, R.; Petherick, E.S.; et al. Cohort Profile: The Born in Bradford multi-ethnic family cohort study. *Int. J. Epidemiol.* **2013**, *42*, 978–991. Available online: <https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dys112> (accessed on 23 March 2020). [CrossRef]
 15. Taylor, K.; McBride, N.; Goulding, N.J.; Burrows, K.; Mason, D.; Pembrey, L.; Yang, T.; Azad, R.; Wright, J.; Lawlor, D.A.; et al. Metabolomics datasets in the Born in Bradford cohort. *Wellcome Open Res.* **2021**, *5*, 264. Available online: <https://wellcomeopenresearch.org/articles/5-264/v2> (accessed on 5 October 2021). [CrossRef]
 16. Yang, Q.; Borges, M.C.; Sanderson, E.; Magnus, M.C.; Kilpi, F.; Collings, P.J.; Soares, A.L.; West, J.; Magnus, P.; Wright, J.; et al. Associations of insomnia on pregnancy and perinatal outcomes: Findings from Mendelian randomization and conventional observational studies in up to 356,069 women. *medRxiv* **2021**, *2021*, 21264689. Available online: <http://medrxiv.org/content/early/2021/10/10/2021.10.07.21264689.abstract> (accessed on 21 January 2022).
 17. Brand, J.S.; Gaillard, R.; West, J.; McEachan, R.R.C.; Wright, J.; Voerman, E.; Felix, J.F.; Tilling, K.; Lawlor, D.A. Associations of maternal quitting, reducing, and continuing smoking during pregnancy with longitudinal fetal growth: Findings from Mendelian randomization and parental negative control studies. *PLoS Med.* **2019**, *16*, e1002972. Available online: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6853297/> (accessed on 20 July 2020). [CrossRef]
 18. Glinianaia, S.V.; Rankin, J.; Wright, C. Congenital anomalies in twins: A register-based study. *Hum. Reprod.* **2008**, *23*, 1306–1311. [CrossRef]
 19. Best, K.E.; Rankin, J. Increased risk of congenital heart disease in twins in the North of England between 1998 and 2010. *Heart* **2015**, *101*, 1807–1812. [CrossRef]
 20. Boyd, A.; Golding, J.; Macleod, J.; Lawlor, D.A.; Fraser, A.; Henderson, J.; Molloy, L.; Ness, A.; Ring, S.; Davey Smith, G. Cohort Profile: The ‘Children of the 90s’—The index offspring of the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **2013**, *42*, 111–127. Available online: <https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dys064> (accessed on 3 March 2020). [CrossRef]
 21. Fraser, A.; Macdonald-Wallis, C.; Tilling, K.; Boyd, A.; Golding, J.; Davey Smith, G.; Henderson, J.; Macleod, J.; Molloy, L.; Ness, A.; et al. Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int. J. Epidemiol.* **2013**, *42*, 97–110. [CrossRef]
 22. Northstone, K.; Lewcock, M.; Groom, A.; Boyd, A.; Macleod, J.; Timpson, N.; Wells, N. The Avon Longitudinal Study of Parents and Children (ALSPAC): An update on the enrolled sample of index children in 2019. *Wellcome Open Res.* **2019**, *4*, 51. Available online: <https://wellcomeopenresearch.org/articles/4-51/v1> (accessed on 10 March 2020). [CrossRef] [PubMed]
 23. Magnus, P.; Birke, C.; Vejrup, K.; Haugan, A.; Alsaker, E.; Daltveit, A.K.; Handal, M.; Haugen, M.; Høiseith, G.; Knudsen, G.P.; et al. Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa). *Int. J. Epidemiol.* **2016**, *45*, 382–388. [CrossRef]
 24. Magnus, P.; Irgens, L.M.; Haug, K.; Nystad, W.; Skjaerven, R.; Stoltenberg, C.; MoBa Study Group. Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). *Int. J. Epidemiol.* **2006**, *35*, 1146–1150. [CrossRef] [PubMed]
 25. Sovio, U.; Goulding, N.; McBride, N.; Cook, E.; Gaccioli, F.; Charnock-Jones, D.S.; Lawlor, D.A.; Smith, G.C.S. A maternal serum metabolite ratio predicts fetal growth restriction at term. *Nat. Med.* **2020**, *26*, 348–353. Available online: <https://www.nature.com/articles/s41591-020-0804-9> (accessed on 20 May 2019). [CrossRef]

26. Bishop, C.; Small, N.; Mason, D.; Corry, P.; Wright, J.; Parslow, R.C.; Bittles, A.H.; Sheridan, E. Improving case ascertainment of congenital anomalies: Findings from a prospective birth cohort with detailed primary care record linkage. *BMJPO* **2017**, *1*, e000171. Available online: <http://bmjpaedsopen.bmj.com/lookup/doi/10.1136/bmjpo-2017-000171> (accessed on 4 March 2020). [CrossRef] [PubMed]
27. Taylor, K.; Thomas, R.; Mumme, M.; Golding, J.; Boyd, A.; Northstone, K.; Caputo, M.; Lawlor, D.A. Ascertaining and classifying cases of congenital anomalies in the ALSPAC birth cohort. *Wellcome Open Res.* **2020**, *5*, 231. Available online: <https://wellcomeopenresearch.org/articles/5-231/v1> (accessed on 13 October 2020). [CrossRef] [PubMed]
28. Leirgul, E.; Fomina, T.; Brodwall, K.; Greve, G.; Holmstrøm, H.; Vollset, S.E.; Tell, G.S.; Øyen, N. Birth prevalence of congenital heart defects in Norway 1994–2009—A nationwide study. *Am. Heart J.* **2014**, *168*, 956–964. Available online: <https://linkinghub.elsevier.com/retrieve/pii/S0002870314004980> (accessed on 18 January 2022). [CrossRef]
29. Smith, G.D.; Lawlor, D.A.; Harbord, R.; Timpson, N.; Day, I.; Ebrahim, S. Clustered Environments and Randomized Genes: A Fundamental Distinction between Conventional and Genetic Epidemiology. *PLoS Med.* **2007**, *4*, e352. [CrossRef]
30. Lawlor, D.A.; Richmond, R.; Warrington, N.; McMahan, G.; Smith, G.D.; Bowden, J.; Evans, D.M. Using Mendelian randomization to determine causal effects of maternal pregnancy (intrauterine) exposures on offspring outcomes: Sources of bias and methods for assessing them. *Wellcome Open Res.* **2017**, *2*, 11. Available online: <https://wellcomeopenresearch.org/articles/2-11/v1> (accessed on 11 November 2020). [CrossRef]
31. Rønningen, K.S.; Paltiel, L.; Meltzer, H.M.; Nordhagen, R.; Lie, K.K.; Hovengen, R.; Haugen, M.; Nystad, W.; Magnus, P.; Hoppin, J.A. The biobank of the Norwegian Mother and Child Cohort Study: A resource for the next 100 years. *Eur. J. Epidemiol.* **2006**, *21*, 619–625. [CrossRef]
32. Lotta, L.A.; Pietzner, M.; Stewart, I.D.; Wittemans, L.B.L.; Li, C.; Bonelli, R.; Raffler, J.; Biggs, E.K.; Oliver-Williams, C.; Auyeung, V.P.W.; et al. A cross-platform approach identifies genetic regulators of human metabolism and health. *Nat. Genet.* **2021**, *53*, 54–64. Available online: <http://www.nature.com/articles/s41588-020-00751-5> (accessed on 16 February 2021). [CrossRef]
33. Hemani, G.; Zheng, J.; Elsworth, B.; Wade, K.H.; Haberland, V.; Baird, D.; Laurin, C.; Burgess, S.; Bowden, J.; Langdon, R.; et al. The MR-Base platform supports systematic causal inference across the human phenome. *elife* **2018**, *7*, e34408. [CrossRef] [PubMed]
34. Day, N.; Oakes, S.; Luben, R.; Khaw, K.T.; Bingham, S.; Welch, A.; Wareham, N. EPIC-Norfolk: Study design and characteristics of the cohort. *European Prospective Investigation of Cancer. Br. J. Cancer* **1999**, *80* (Suppl. S1), 95–103. [PubMed]
35. Di Angelantonio, E.; Thompson, S.G.; Kaptoge, S.; Moore, C.; Walker, M.; Armitage, J.; Ouweland, W.H.; Roberts, D.J.; Danesh, J.; INTERVAL Trial Group. Efficiency and safety of varying the frequency of whole blood donation (INTERVAL): A randomised trial of 45,000 donors. *Lancet* **2017**, *390*, 2360–2371. [CrossRef]
36. Myers, T.A.; Chanock, S.J.; Machiela, M.J. LDlinkR: An R Package for Rapidly Calculating Linkage Disequilibrium Statistics in Diverse Populations. *Front. Genet.* **2020**, *11*, 157. [CrossRef]
37. Taylor, K. Using an Untargeted Metabolomics Platform to Explore Associations between Maternal Metabolites and Congenital Heart Disease in the Offspring. 2020. Available online: <https://osf.io/zf4k6/> (accessed on 16 February 2021).
38. McBride, N.; Yousefi, P.; Sovio, U.; Taylor, K.; Vafai, Y.; Yang, T.; Hou, B.; Suderman, M.; Relton, C.; Smith, G.C.; et al. Do mass-spectrometry-derived metabolomics improve prediction of pregnancy-related disorders? Findings from a UK birth cohort with independent validation. *medRxiv* **2021**, *2021*, 21256218. Available online: <http://medrxiv.org/content/early/2021/05/04/2021.05.04.21256218.abstract> (accessed on 1 January 2021). [CrossRef]
39. Lai, G.; Li, X.; Zhang, B.; Wang, L.; He, R.; Zhao, Y. Decreased Amino Acid Concentrations are Involved in Congenital Heart Disease. *Ann. Nutr. Metab.* **2019**, *74*, 257–263. [CrossRef]
40. Parsons, A.M.; Bouma, G.J. A Potential Role and Contribution of Androgens in Placental Development and Pregnancy. *Life* **2021**, *11*, 644. Available online: <https://www.mdpi.com/2075-1729/11/7/644> (accessed on 6 January 2022). [CrossRef]
41. Mai, M.; Tönjes, A.; Kovacs, P.; Stumvoll, M.; Fiedler, G.M.; Leichtle, A.B. Serum Levels of Acylcarnitines Are Altered in Prediabetic Conditions. *PLoS ONE* **2013**, *8*, e82459. [CrossRef]
42. Innis, S.M. Fatty acids and early human development. *Early Hum. Dev.* **2007**, *83*, 761–766. [CrossRef] [PubMed]
43. McKeegan, P.J.; Sturme, R.G. The role of fatty acids in oocyte and early embryo development. *Reprod. Fertil. Dev.* **2011**, *24*, 59–67. [CrossRef]
44. Gittenberger-de Groot, A.C.; Bartelings, M.M.; Deruiter, M.C.; Poelmann, R.E. Basics of Cardiac Development for the Understanding of Congenital Heart Malformations. *Pediatr. Res.* **2005**, *57*, 169–176. Available online: <http://www.nature.com/doi/10.1203/01.PDR.0000148710.69159.61> (accessed on 4 September 2020). [CrossRef] [PubMed]
45. Mills, H.L.; White, S.L.; Pasupathy, D.; Briley, A.L.; Santos Ferreira, D.L.; Seed, P.T.; Nelson, S.M.; Sattar, N.; Tilling, K.; Poston, L.; et al. The effect of a lifestyle intervention in obese pregnant women on gestational metabolic profiles: Findings from the UK Pregnancies Better Eating and Activity Trial (UPBEAT) randomised controlled trial. *BMC Med.* **2019**, *17*, 15. Available online: <https://bmcmecine.biomedcentral.com/articles/10.1186/s12916-018-1248-7> (accessed on 26 March 2020). [CrossRef] [PubMed]
46. Sanderson, E.; Richardson, T.G.; Morris, T.T.; Tilling, K.; Davey Smith, G. Estimation of causal effects of a time-varying exposure at multiple time points through Multivariable Mendelian randomization. *medRxiv* **2022**, *2022*, 22268740. Available online: <http://medrxiv.org/content/early/2022/01/05/2022.01.04.22268740.abstract> (accessed on 21 January 2022). [CrossRef]
47. Yang, J.; Kang, Y.; Cheng, Y.; Zeng, L.; Yan, H.; Dang, S. Maternal Dietary Patterns during Pregnancy and Congenital Heart Defects: A Case-Control Study. *Int. J. Environ. Res. Public Health* **2019**, *16*, 2957. [CrossRef]