

Published in final edited form as:

Early Hum Dev. 2014 November ; 90(11): 769–780. doi:10.1016/j.earlhumdev.2014.08.023.

Approaches for drawing causal inferences from epidemiological birth cohorts: A Review

Rebecca C Richmond, BA(Hons)¹, Aleef Al-Amin², George Davey Smith, MA, MD, BChir, DSc³, and Caroline L Relton, BSc, PhD^{4,*}

Rebecca C Richmond: rebecca.richmond@bristol.ac.uk; Aleef Al-Amin: aleefalamin@gmail.com; George Davey Smith: kz.davey-smith@bristol.ac.uk

¹Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, UK

²University of Bristol Medical School, University of Bristol, Bristol, UK

³Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, UK, Telephone: +44 (0)117 3310099 Fax: +44 (0)117 3310123

⁴Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, UK, Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK, Telephone: +44 (0)117 3314028, Fax: +44 (0)117 3310123

Abstract

Large-scale population-based birth cohorts, which recruit women during pregnancy or at birth and follow up their offspring through infancy and into childhood and adolescence, provide the opportunity to monitor and model early life exposures in relation to developmental characteristics and later life outcomes. However, due to confounding and other limitations, identification of causal risk factors has proved challenging and published findings are often not reproducible. A suite of methods has been developed in recent years to minimise problems afflicting observational epidemiology, to strengthen causal inference and to provide greater insights into modifiable intra-uterine and early life risk factors. The aim of this review is to describe these causal inference methods and to suggest how they may be applied in the context of birth cohorts and extended along with the development of birth cohort consortia and expansion of “omic” technologies.

Keywords

birth cohort; causal inference; consortia; DOHaD; epidemiology; epigenetics; life course; metabolomics; omics

*Correspondence: caroline.relton@bristol.ac.uk.

Conflicts of interest statement

None declared.

Introduction

Large-scale population-based birth cohorts recruit women during pregnancy or at birth over a defined time period and follow up their offspring through infancy and into childhood and adolescence. The longitudinal design of these cohorts is a key feature, providing the opportunity to monitor and model early life exposures in relation to developmental characteristics and later life outcomes, with prospective data collected at repeat follow-ups. Data are often collected on both parents and offspring and include information on demographic, socio-economic and lifestyle characteristics and environmental exposures obtained from questionnaires, clinic data for assessing health and development, and data from biological samples. Some cohorts have been designed as multipurpose resources, whilst others focus on specific health or exposure-related research questions. The size of birth cohorts varies considerably, from a few hundred individuals to over 100,000 in countries where population-based record linkage is possible.

A major focus of such studies is exposure to risk factors during early life developmental periods which can have important consequences for health and disease. The “Developmental Origins of Health and Disease” (DOHaD) hypothesis outlines how the risk of chronic disease in adult life is initially induced through biological programming of the foetus or infant in response to early environmental signals 1,2. These responses include molecular, hormonal, metabolic or physiological changes which may have negative impacts on later health. Of particular interest is the data captured on maternal exposures acting during pregnancy, driven by the notion that the intra-uterine environment is a critical period for influencing offspring development and programming events 3,4.

Studies have reported associations between foetal growth, maternal nutrition, exposure to drugs, pollutants and hormones in-utero and a whole host of perinatal and later life offspring traits. The influence of postnatal factors has also been explored, including early life growth 5 and breastfeeding 6. Of particular value are historical birth cohorts which can be used to study the influence of early life exposures on later disease 7. As well as DOHaD, other aspects of research within lifecourse epidemiology may be investigated within the context of a birth cohort 8,9 and details of these can be found elsewhere 10.

An attractive feature of birth cohorts is the ability to obtain information on other family members, not only the mothers of the offspring, but sometimes fathers, siblings and grandparents. Family-based sampling can facilitate inter-generational studies of the influence of parental characteristics on a range of offspring outcomes and may aid in disentangling the genetic determinants of disease from environmental risk factors 11.

Increasingly, birth cohorts collect and store biosamples from their participants, which can be used to obtain genetic, epigenetic and metabolic profiles, and to measure biomarkers of environmental exposures such as smoking and pollutants. Biosampling allows the exploration of how social and environmental factors leave biological imprints, independent of or in combination with genetic background. The ‘omics’ revolution 12 offers the potential to explore putative mechanisms by which specific exposures convey disease risk, whereby

identified molecules provide robust biomarkers of early life exposure or may act as intermediates in pathways between exposure and risk of later outcomes.

In addition to the wealth of data collected, longitudinal birth cohorts can offer more to observational epidemiology than other study designs because they allow for prospective time-ordering of the associations of interest i.e. with exposures preceding outcomes, which is useful for establishing causality. However, a key limitation to causal inference in epidemiological birth cohorts is potential confounding, leading to spurious observational associations 13,14. Distinguishing causality from correlation is essential to identify key early life modifiable causes of ill health and disease and to uncover new mechanistic pathways for therapeutic intervention. A suite of methods has been developed in the last decade to minimise problems afflicting observational epidemiology and to strengthen causal inference. The aim of this review is to describe the causal inference methods that have been used to provide greater insights into modifiable intra-uterine and early life risk factors in the context of large epidemiological birth cohorts and to suggest how we may improve methodological approaches, especially in relation to the expansion of “omics” technologies.

Challenges of establishing causality in birth cohorts

Key problems of observational epidemiology which limit its ability to establish causal effects include: 1) reverse causation - where the outcome of interest affects the exposure; 2) confounding – the presence of common causes of the risk factor of interest and the outcome; 3) selection bias – when the study participants are selected in a manner that biases the effect estimate in an association; 4) measurement error in the exposure, confounding factors or outcome. The characteristics of birth cohorts are such that some of these problems can be minimised. For example, their prospective study design means that there is no biased retrospective assessment and the likelihood of reverse causation is reduced due to the time-ordering of the exposure-outcome associations. These studies also allow for repeated measures to be taken at different time points and appropriate analytical techniques may be used to account for missing data, reducing the role of measurement error and selection bias 15,16.

Observational epidemiology undertaken in the context of a birth cohort generally relies on the assumption that all confounding characteristics have been identified and measured with little or no error. However, confounders may be inadequately measured (residual confounding) or there may be unobserved factors (unmeasured confounding) 17 which can lead to spurious associations and conclusions about intra-uterine and early life risk factors 18,19. Inconsistent findings between cohorts and randomized controlled trials (RCTs) highlight the methodological challenges in establishing robust causal links 13,20. For example, in observational studies maternal vitamin C intake has been found to be associated with higher birth weight in the offspring 21. However, large RCTs where pregnant women have been randomized to vitamin C supplements 22–24 have found no benefit of supplementation on birth weight. These conflicting findings are likely due to confounding in the observational association, as mothers with higher vitamin C intake tend to have lower rates of smoking and are from a higher socioeconomic background, which influence birth weight 25.

Other limitations introduced by the very nature of birth cohorts include the long time gap between outcomes and exposures, increasing the likelihood of confounding. Another implication of this time gap is the relevance of early life exposures experienced when the birth cohorts were established to contemporary cohorts. Finally, given the high correlation between maternal exposures and behaviours in pregnancy with those postnatally it is often difficult to tease apart intra-uterine from postnatal effects 26.

Classic epidemiological approaches for drawing causal inferences

Data collected on parents, offspring and other family members in epidemiological birth cohorts may be integrated in a suite of methods which minimise problems of confounding, strengthen causal inference and provide greater insights into modifiable early life risk factors. Strength of evidence obtained from these methods can be placed between observational associations and RCTs in the hierarchy of evidence for clinical guideline production. Table 1 includes a selection of large, well-established cohorts and the data available in these cohorts which may permit the application of the causal inference methods described in this review. Table 2 outlines each of the main causal inference methods, with examples and linked schematic diagrams in Figure 1.

Randomized Controlled Trials

Well-conducted, large RCTs are the gold standard for estimating causal effects in population health and this is also the case in the setting of early life influences, for example with the randomization of women to different interventions in pregnancy. A number of RCTs of pregnancy and early life interventions originally set up to investigate short-term outcomes have been extended to follow up offspring at multiple ages. One example of a birth cohort nested within in an RCT is the PROBIT trial 27,28. This cluster-randomized controlled trial involved randomization to a breastfeeding promotion intervention which resulted in longer duration of any and exclusive breastfeeding and has been used to investigate the causal effect of breastfeeding on later health outcomes, including obesity, blood pressure, cognitive function and eating attitudes 29–33. RCTs require large investment and their experimental nature means that they should be reserved for interventions that have strong support from observational epidemiology. In addition, for some exposures it is not possible or would be unethical to randomize participants and where RCTs are conducted, they are often done so in selected populations and so findings may not be generalizable.

Cross – cohort comparisons

Support for the initiation of the PROBIT trial came from observational studies which have shown breastfeeding to be protective against a wide range of later outcomes. However, not all of these associations persist in a randomized trial setting 29–31. This discordance can be explained by the fact that the majority of observational studies have been conducted in higher-income countries where breastfeeding is strongly related to higher socio-economic circumstances, maternal non-smoking and healthy diet. The links between breastfeeding and these factors would generate non-causal observational associations between breastfeeding and health outcomes, and the ability to fully evaluate and statistically adjust for such

confounding is limited. One way to circumvent this problem, without initiating an RCT, would be to compare associations between two or more populations in which the underlying confounding structures are markedly different. For example, if the associations found in higher-income countries are causal then one would expect them to be found in lower-middle income countries where breastfeeding is often not associated with socio-economic position 34. An analysis of a UK-based cohort study, ALSPAC, and a Brazilian-based cohort study, Pelotas, showed that the inverse association of breastfeeding with later offspring body mass index (BMI) and blood pressure found in higher income countries is not present in lower middle-income countries. By contrast, a positive association with intelligence quotient (IQ) was found in both settings 34. These findings have been validated by results of the PROBIT study, based in the middle-income country of Belarus 27,28. The assumption about different confounding structures in different cohorts may not be correct and has to be thoroughly investigated. In addition, harmonisation of variables between cohorts is required in order to minimise the influence of statistical heterogeneity.

Negative controls

It is also possible to infer a causal effect by comparing an observed association between a particular exposure and an outcome with a negative control. A negative control situation is one that cannot involve the hypothesised causal mechanism, but which is likely to involve the same sources of bias or confounding as in the original association 18,35,36. For example, the association of an exposure and outcome may be compared with that of another exposure, which is equally socially patterned, and the same outcome. A study conducted in the Norwegian Mother and Child cohort (MoBa) compared the magnitude of association between maternal folic acid supplementation in pregnancy and children's risk of autistic disorders with the association between maternal fish oil supplementation and autistic disorders. A reduced risk of autistic disorder in children of folic acid users was evident but no such association was found with prenatal fish oil use, even though fish oil use was associated with similar socio-economic characteristics as folic acid use 37.

Negative controls can also be used if one wishes to investigate whether an association between a particular exposure and outcome arises in a proposed critical period, such as in-utero. For example, maternal smoking after pregnancy would not be expected to have the same influence on offspring outcomes as smoking during pregnancy if the mechanism of influence is through the intra-uterine environment 38. However, the high correlation of pre- and postnatal smoking makes it difficult to disentangle causal effects 26 and women who do not smoke in pregnancy but do postnatally may be characteristically different from women who continue to smoke, which may re-introduce confounding. This can be avoided through the use of within-individual comparisons or when the influence on exposure patterns is externally generated 18,19. For example, the Dutch Hunger Winter study demonstrates the specific effect of imposed nutritional deprivation during early pregnancy on a number of health outcomes, compared with women who experienced famine at other stages in pregnancy 39.

Parental comparisons

A negative control design primarily used for exploring the extent to which associations of intra-uterine exposure might be causally related to offspring outcomes in later life is the parental comparisons approach. If there is a causal intra-uterine effect, one would expect a stronger maternal-offspring association than paternal-offspring association for the same exposure assessed at the time of pregnancy. Where associations are similar for both parents it is likely that there is confounding by genetic or shared environmental characteristics 11,18,36. Proof of concept has been illustrated with maternal smoking in pregnancy which is strongly associated with lower offspring birth weight, whereas paternal smoking is only weakly associated. When both maternal and paternal smoking during pregnancy are taken into account, the former association is little attenuated whereas the latter association is essentially abolished, arguing for a biological effect of maternal smoking in pregnancy on offspring birth weight 18.

It has been hypothesised that maternal obesity and metabolic profiles related to this may, during pregnancy, programme the offspring for greater risk of obesity in later life 40,41. This could result in intergenerational acceleration, with ever-increasing levels of obesity in the population 42. Some parental comparison studies find stronger associations of maternal BMI than paternal BMI with offspring BMI 43–45, although these have often been of small sample size, with different sources and degrees of validity for BMI measures, and non-paternity for biological measures has generally not been taken into account 46. Subsequent studies addressing these issues have found that maternal and paternal BMI relate very similarly to offspring adiposity 46–50, arguing against a major specific effect of the intra-uterine environment and suggesting that the associations are driven by shared familial genetic or lifestyle characteristics.

Some evidence has been found which supports potential male-line transgenerational responses, invoking parent of origin, imprinting and epigenetic phenomena 51,52. Maternal and paternal associations of similar magnitude may therefore be interpreted as showing intra-uterine maternal influences which are offset by these paternal pathways. However, it has been posited that the likelihood of such perfectly matched effects being produced by mechanistically distinct processes is low 18,53.

Sibling comparisons

It may be possible to compare outcomes within siblings who are concordant or discordant for early life exposures. Since familial background will generally be similar for siblings, comparing outcome differences in relation to discordant exposures within sibships effectively “matches” on family characteristics, providing a stronger means of controlling for certain confounding factors 11. Such study designs have been used to show that gestational diabetes 54,55, gestational weight gain 56 and extreme BMI 57,58 are likely to be causally related to later offspring obesity and other metabolic outcomes 41, with findings being translated into long-term follow-up of participants in randomized controlled trials 59,60.

Again there are instances where this causal analysis method has provided contrasting results in different studies. For example, sibling studies have been used to explore whether the positive association between birth weight and later IQ 61 is causal. While some studies suggest that birth weight differences within sibships are related to differences in intelligence, implying an intrauterine effect 62,63, others show no evidence of association 64,65, arguing that the association observed in the population may be explained by factors such as family socioeconomic background.

It is important to bear in mind that, although sibling comparison estimates will not be influenced by unmeasured familial confounders, there are notable limitations to this study design which may explain the discrepancy in findings 66. Such estimates are more severely biased by non-shared confounders than population-level comparisons 67 and are more sensitive to misclassification of the exposure and measurement error 66,68. Use of a sibling comparison design also limits the population included, affecting power and demonstrating the need for large sample sizes to obtain robust causal evidence.

Mendelian randomization

Mendelian randomization (MR) is a method that utilises genetic variants robustly associated with modifiable exposures to infer causality 69. The MR design is analogous to an RCT, where study participants are randomly allocated to a treatment to avoid potential confounding between treatment and outcome 70. MR creates a similar scenario by exploiting Mendel's laws (segregation and independent assortment). Given these laws, at a population level genetic variants should not be associated with genetic or environmental confounding factors that can distort conventional observational studies. Analysing data according to genotype will therefore compare groups that differ by an on-average level of a modifiable exposure, but not by a myriad of behavioural, social and physiological variables that may confound observational associations 71,72. In addition, in a genetic association the direction of causation is from genetic variation to the outcome, and not vice versa as disease processes do not alter germline genotype. Genetic variants are also subject to relatively little measurement error or bias and variants will generally be related to a modifiable exposure throughout life, avoiding attenuation by errors 73.

Where maternal genotype is taken to be a proxy for environmentally-modifiable exposures in pregnancy, this may provide unique insights into the causal nature of intra-uterine environment influences on later offspring outcomes 18. For example, variation in *MTHFR* is associated with methyltetrahydrofolate reductase activity and hence with circulating folate and homocysteine levels. Maternal *MTHFR* variants have been found to influence risk of neural tube defects (NTD) in offspring 74, implying a causal effect of low maternal folate. These findings are consistent with the results of RCTs of maternal folate supplementation which is associated with reduced risk of offspring congenital abnormalities 75,76. In this example, the effect of maternal genotype on risk of NTD was greater than paternal or offspring genetic estimates, implying an independent maternal effect 74 which is consistent with the hypothesis that maternal folate intake is the exposure of importance.

Limitations of the Mendelian randomization approach have been outlined in detail elsewhere 77,78, and include low statistical power due to the small amount of variance in a trait explained by the genetic variant; population stratification, which may induce confounding when allele frequencies and disease risk differ according to the genetic ancestry of populations within the study; and pleiotropy, where the genetic variant influences more than one post-transcriptional process and may affect the outcome via a pathway that is independent of the exposure. Methods may be implemented to address these limitations and extensions of the MR approach applied to avoid them 77,78.

Non-genetic instrumental variable analysis

The use of genotype in MR studies is an application of instrumental variable (IV) analysis 79,80, which may be used to obtain an estimate for the magnitude of a causal effect. An IV is a variable that is associated with the outcome only through its robust association with the exposure, and therefore an IV will typically not be associated with factors that confound the association of exposure and outcome. Examples of non-genetic instrumental variables include external factors which influence a population largely at random, such as the famine experienced in the Dutch Hunger Winter 39, climate conditions 81, or cigarette taxation 82. However, in these cases the external or “exogenous” factor is generally rare or of small effect. Another non-genetic IV which is more commonplace is the phenotype of a family member in family-based studies, which may be used to proxy for own phenotype. For example, offspring anthropometry has been used as an IV for examining the causal effect of own anthropometry on mortality 83,84. As offspring anthropometry is likely influenced by the same socio-economic, lifestyle and genetic confounders as parental anthropometry, this method is used primarily to deal with reverse causation, under the assumption that offspring’s anthropometry will not be influenced by parent’s illness.

Triangulation

The above causal inference methods have different underlying assumptions, strengths and limitations and an integration of different approaches to the same research question may be used to improve the identification and estimation of causal effects through the “triangulation” of findings. This may be done under the supposition that independent biases are unlikely to lead to the same result across a range of methodological approaches. If causal effects are consistently estimated, the likelihood that they are unbiased is high. If they differ between the approaches, there is a further need to investigate whether the underlying assumptions for each approach have been violated. One example of triangulation has already been alluded to, which is the similarity in findings between a cross-cohort comparison study 34 and a randomized controlled trial investigating the effect of breastfeeding on offspring BMI, blood pressure and IQ 29–32. Conventional multiple regression, parental comparison, between-sibling analyses, Mendelian randomization, non-genetic instrumental variable and RCT studies have all been consistent in their findings of a causal effect of maternal smoking in pregnancy on offspring birth weight 85. “Triangulation” methods have also been exemplified within single studies where two complementary approaches have shown consensus on early life causal effects 45,54. The approach of privileging a hypothesis which

fits with the overall pattern of findings and knowledge across all informative sources is within the tradition of “inference to the best explanation” approaches to causal reasoning 86.

Consortia

One characteristic which all of the described causal inference methods have in common is that they are often underpowered and generally require large sample sizes. Therefore, as well as using triangulation, there is a need for independent replication of findings in order to avoid spurious conclusions in causal inference analysis. Cross-cohort analysis can improve power and statistical precision, and can provide high quality evidence on the causal effects of early life exposures on later health and disease. Collaboration is already evident in some instances, with the pooling and harmonising of data to address research questions on environmental exposures 87,88 and genetic associations 89–91. There are several examples of birth cohort collaborations, including CHICOS (<http://www.chicosproject.eu/the-project/management/>), EAGLE (<http://www.copsac.com/content/eagle-consortium>), EGG (<http://egg-consortium.org/>) and ENRIECO (<http://www.enrieco.org/>) 92 and a tool for accessing information on each birth cohort has been made available at <http://www.birthcohorts.net/> 93. Also of importance in this field is the inclusion of birth cohort studies from low- and middle-income countries 94,95, where variation in environmental exposures, health outcomes and confounding structures may be used to improve causal inference 34. To date, collaborations have been used to replicate findings from causal inference analysis in multiple cohorts, including parental comparisons 96 and Mendelian randomization 97.

New data

An attribute of many birth cohorts is their biological sampling which includes the collection of blood, urine and hair samples. New technologies permit genotyping and profiling of methylation, metabolites and biomarkers of environmental exposures, and open up new avenues for exploring underlying causal pathways. Of particular value is the collection of serial samples from the same individuals in some birth cohorts, which allows assessment of change in molecular measures over time.

As shown in Table 1, many birth cohorts now have genome-wide data available on a large number of individuals, including both offspring and parents. These may be used in Genome Wide Association Studies (GWAS), where associations between a wide range of phenotypes and genetic variants across the genome are determined in a hypothesis-free approach. More recently, an innovative method utilising genome-wide data in mothers and offspring has been developed which allows the delineation of maternal-specific influences on offspring outcomes⁹⁸. The ability to identify many robust genotype-phenotype associations is of merit for Mendelian randomization analysis which has classically involved the use of a single variant to proxy for a particular modifiable exposure. GWAS has uncovered a host of genetic variants which explain an increasing proportion of the variance in a trait and may act as a stronger instrument for improving the precision of causal estimates⁹⁹. The use of genetic scores, created by adding up the total number of risk alleles a person has, offers particular promise in this regard^{100,101}. However, as the function of a variant identified in GWAS is often unknown, the assumption that it will only influence the outcome through its

direct effect on the exposure is difficult to assert. Nonetheless strategies exist for assessing potential pleiotropy 72,99.

Building on the success of GWAS and the availability of cost effective and robust technologies is the use of “omics” within population health science. This is largely concerned with understanding how gene regulatory mechanisms or gene products interact with the environment to influence health-related outcomes and is useful for investigating the molecular pathways that may underpin causal effects. Of particular utility are large-scale epigenetic and metabolomic scans for formulating novel hypotheses on biological processes. However, in contrast to germ-line genetic variation, epigenetic and metabolomic signatures are largely phenotypic, and are subject to the same problems of confounding and reverse causation which afflict conventional epidemiology 53,102,103 (Figure 2). The extension of causal inference approaches is therefore of particular utility in determining causal associations between “omic” markers and a range of exposures and outcomes 77.

Epigenetics

Epigenetic mechanisms are involved in regulating gene activity which creates phenotypic variation without altering the underlying DNA code. Epigenetics is a potentially major mechanism by which environmental factors can affect physiological function and disease risk. In particular, DNA methylation has become increasingly integrated into population-based studies as a potential modifiable indicator of the underlying biological changes.

Epidemiological approaches can be used to identify whether epigenetic processes are involved in mediating the association between various risk factors and common complex disease 104,105. Longitudinal cohort studies that make use of multiple time points are useful for investigating how the epigenome changes over time, as a result of varying exposures, and how this contributes to disease development 106. In particular, there is considerable interest in the role of epigenetic mechanisms in DOHaD as epigenetic states are often established in early development 107–109. This makes birth cohorts with sample collection from pregnant women and offspring at birth of particular value for providing insights into the temporal relationship between early life exposures and epigenetic changes 110–112, which may then predict later health-related outcomes 113–115.

It is important to bear in mind that epigenetic profiles can be influenced by technical or genetic factors, cellular and tissue heterogeneity, time-varying artefacts and stochastic changes. These sources of noise threaten the detection of biological signals and the ability to infer causality from associations 53,103. Careful study design, data collection and control of sources of variability are therefore required, as are methods which will contribute to the identification of predictive epigenetic biomarkers and modifiable targets for intervention 102,116,117.

Many of the approaches already listed to address causality in conventional epidemiological settings can also be used to interrogate causality in associations involving epigenetic changes. For example, maternal smoking in pregnancy has been shown to be associated with DNA methylation in newborns 118 and the finding of no paternal associations highlights the

prominent intra-uterine influence of maternal smoking on offspring DNA methylation at birth 119 and at later ages in the offspring 120. Mendelian randomization analysis has also been used in the context of epigenetic epidemiology to investigate the causal effect of maternal red blood cell folate on genome-wide methylation in infant cord blood 121, using the previously described *MTHFR* genotype as an instrument. However, further work is needed to investigate whether the identified methylation changes mediate the influence of intra-uterine exposures on developmental outcomes, for example in a “two-step Mendelian randomization” framework 77,102,116,117.

Metabolomics

Metabolomics is an emerging technology involving the measurement of metabolites which likely act as intermediates in biological pathways. An advantage of using metabolites as intermediate phenotypes is that they are more proximal to biological pathways than downstream phenotypes or clinical endpoints 122, boosting the statistical power to detect associations 123,124. Metabolites are also useful in birth cohorts when disease endpoints have not yet been reached.

However, as metabolites are influenced by both genetic and environmental factors and by disease processes, they too are prone to the limitations of observational study. Once an association between a metabolite and a trait has been observed, the next challenge is to distinguish causal effects, with potential implications for clinical outcomes and disease pathogenesis, from non-causal associations, which may have potential implications for biomarker discovery 12,125. Different statistical methodologies may be used to construct a causal framework involving metabolites, and to dissect causal relationships 126. This framework also suggests the usefulness of “triangulating” causal inference methods in the domain of high-dimensional molecular data as an exploratory tool to infer causal relationships.

Summary

This review has outlined a suite of causal inference methods including cross-cohort comparisons, negative control studies, sibling studies, Mendelian randomization analysis and instrumental variable techniques. These methods make use of the wide range of data available in epidemiological birth cohorts in order to establish causal links between early life influences and a range of developmental and health outcomes. Such methods have often been shown to produce the same conclusions regarding causal effects as randomized controlled trials, which are not always feasible or ethical, and may be used to inform on interventions. Strengthening causal inference is also an important step in “omics” research for distinguishing causal molecular pathways that may underpin causal effects of early life exposures on complex traits and diseases.

The methods for causal inference described enhance capability to interpret conventional observational associations, though some discrepancies in findings between studies highlight their limitations, in particular their lack of power in small samples. An integration or “triangulation” of different approaches to the same research question may be used to

improve the identification and estimation of causal effects in observational data. In addition, cross-cohort analysis and the independent replication of findings can improve power and statistical precision and provide more high-quality evidence for causality. This may be enabled with collaboration among different birth cohorts and the dissemination and harmonisation of techniques through the established consortia.

Acknowledgements

RCR, GDS and CLR are members of the MRC Integrative Epidemiology Unit (IEU) funded by the UK Medical Research Council (MC_UU_12013) and the University of Bristol. RCR is funded by a Wellcome Trust 4-year PhD studentship (Grant Code: WT083431MF). GDS and CLR are partially supported by the ESRC (RES-060-23-0011) "The biosocial archive: transforming lifecourse social research through the incorporation of epigenetic measures". GDS's work is supported in part by the European Research Council grant DEVHEALTH 269874.

References

1. Gillman MW. Developmental origins of health and disease. *The New England Journal of Medicine*. 2005; 353:1848–50. [PubMed: 16251542]
2. Gluckman, P.; Hanson, Me. *The developmental origins of health and disease*. Cambridge: Cambridge University Press; 2006.
3. Barker DJ. The fetal and infant origins of adult disease. *BMJ*. 1990; 301:1111. [PubMed: 2252919]
4. Barker, DJP. *Mothers, babies and disease in later life*. London: Churchill Livingstone; 1994.
5. Ong KK, Ahmed ML, Emmett PM, Preece MA, Dunger DB. Association between postnatal catch-up growth and obesity in childhood: prospective cohort study. *BMJ*. 2000; 320:967–71. [PubMed: 10753147]
6. Bergmann KE, Bergmann RL, Von Kries R, et al. Early determinants of childhood overweight and adiposity in a birth cohort study: role of breast-feeding. *International journal of obesity and related metabolic disorders : Journal of the International Association for the Study of Obesity*. 2003; 27:162–72.
7. Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). *International Journal of Epidemiology*. 2006; 35:34–41. [PubMed: 16155052]
8. Ben-Shlomo Y, Kuh D. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *International Journal of Epidemiology*. 2002; 31:285–93. [PubMed: 11980781]
9. Kuh D, Ben-Shlomo Y, Lynch J, Hallqvist J, Power C. Life course epidemiology. *Journal of Epidemiology and Community Health*. 2003; 57:778–83. [PubMed: 14573579]
10. Power C, Kuh D, Morton S. From developmental origins of adult disease to life course research on adult disease and aging: insights from birth cohort studies. *Annual Review of Public Health*. 2013; 34:7–28.
11. Lawlor, DA.; Mishra, Ge. *Family Matters. Designing, analyzing and understanding family-based studies in life course epidemiology*. Oxford: Oxford University Press; 2009.
12. Vineis P, van Veldhoven K, Chadeau-Hyam M, Athersuch TJ. Advancing the application of omics-based biomarkers in environmental epidemiology. *Environmental and Molecular Mutagenesis*. 2013; 54:461–7. [PubMed: 23519765]
13. Davey Smith G. Reflections on the limitations to epidemiology. *Journal of Clinical Epidemiology*. 2001; 54:325–31. [PubMed: 11297881]
14. Davey Smith G, Phillips AN. Confounding in epidemiological studies: why "independent" effects may not be all they seem. *BMJ*. 1992; 305:757–9. [PubMed: 1298234]
15. Spratt M, Carpenter J, Sterne JA, et al. Strategies for multiple imputation in longitudinal studies. *American Journal of Epidemiology*. 2010; 172:478–87. [PubMed: 20616200]
16. Tilling K, Davies NM, Nicoli E, et al. Associations of growth trajectories in infancy and early childhood with later childhood outcomes. *The American Journal of Clinical Nutrition*. 2011; 94:1808S–13S. [PubMed: 21633072]

17. Fewell Z, Davey Smith G, Sterne JA. The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. *American Journal of Epidemiology*. 2007; 166:646–55. [PubMed: 17615092]
18. Davey Smith G. Assessing intrauterine influences on offspring health outcomes: can epidemiological studies yield robust findings? *Basic & Clinical Pharmacology & Toxicology*. 2008; 102:245–56. [PubMed: 18226080]
19. Davey Smith G, Leary S, Ness A, Lawlor DA. Challenges and novel approaches in the epidemiological study of early life influences on later disease. *Advances in Experimental Medicine and Biology*. 2009; 646:1–14. [PubMed: 19536658]
20. Lawlor DA, Davey Smith G, Kundu D, Bruckdorfer KR, Ebrahim S. Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? *Lancet*. 2004; 363:1724–7. [PubMed: 15158637]
21. Mathews F, Yudkin P, Neil A. Influence of maternal nutrition on outcome of pregnancy: prospective cohort study. *BMJ*. 1999; 319:339–43. [PubMed: 10435950]
22. Poston L, Briley AL, Seed PT, Kelly FJ, Shennan AH, Vitamins in Pre-eclampsia Trial C. Vitamin C and vitamin E in pregnant women at risk for pre-eclampsia (VIP trial): randomised placebo-controlled trial. *Lancet*. 2006; 367:1145–54. [PubMed: 16616557]
23. Villar J, Purwar M, Meriardi M, et al. World Health Organisation multicentre randomised trial of supplementation with vitamins C and E among pregnant women at high risk for pre-eclampsia in populations of low nutritional status from developing countries. *BJOG : International Journal of Obstetrics and Gynaecology*. 2009; 116:780–8.
24. Roberts JM, Myatt L, Spong CY, et al. Vitamins C and E to prevent complications of pregnancy-associated hypertension. *New England Journal of Medicine*. 2010; 362:1282–91. [PubMed: 20375405]
25. Robinson SM, Crozier SR, Borland SE, Hammond J, Barker DJ, Inskip HM. Impact of educational attainment on the quality of young women's diets. *European Journal of Clinical Nutrition*. 2004; 58:1174–80. [PubMed: 15054431]
26. Mishra G, Nitsch D, Black S, De Stavola B, Kuh D, Hardy R. A structured approach to modelling the effects of binary exposure variables over the life course. *International Journal of Epidemiology*. 2009; 38:528–37. [PubMed: 19028777]
27. Kramer MS, Chalmers B, Hodnett ED, et al. Promotion of Breastfeeding Intervention Trial (PROBIT): a randomized trial in the Republic of Belarus. *JAMA*. 2001; 285:413–20. [PubMed: 11242425]
28. Patel R, Oken E, Bogdanovich N, et al. Cohort profile: The promotion of breastfeeding intervention trial (PROBIT). *International Journal of Epidemiology*. 2014; 43:679–90. [PubMed: 23471837]
29. Kramer MS, Matush L, Vanilovich I, et al. Effects of prolonged and exclusive breastfeeding on child height, weight, adiposity, and blood pressure at age 6.5 y: evidence from a large randomized trial. *American Journal of Clinical Nutrition*. 2007; 86:1717–21. [PubMed: 18065591]
30. Martin RM, Patel R, Kramer MS, et al. Effects of promoting longer-term and exclusive breastfeeding on adiposity and insulin-like growth factor-I at age 11.5 years: a randomized trial. *JAMA*. 2013; 309:1005–13. [PubMed: 23483175]
31. Martin RM, Patel R, Kramer MS, et al. Effects of promoting longer-term and exclusive breastfeeding on cardiometabolic risk factors at age 11.5 years: a cluster-randomized, controlled trial. *Circulation*. 2014; 129:321–9. [PubMed: 24300437]
32. Kramer MS, Aboud F, Mironova E, et al. Breastfeeding and child cognitive development: new evidence from a large randomized trial. *Archives of General Psychiatry*. 2008; 65:578–84. [PubMed: 18458209]
33. Skugarevsky O, Wade KH, Richmond RC, et al. Effects of promoting longer-term and exclusive breastfeeding on childhood eating attitudes: a cluster-randomized trial. *International Journal of Epidemiology*. 2014; 43:1263–71. [PubMed: 24706729]
34. Brion MJ, Lawlor DA, Matijasevich A, et al. What are the causal effects of breastfeeding on IQ, obesity and blood pressure? Evidence from comparing high-income with middle-income cohorts. *International Journal of Epidemiology*. 2011; 40:670–80. [PubMed: 21349903]

35. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology*. 2010; 21:383–8. [PubMed: 20335814]
36. Davey Smith G. Negative control exposures in epidemiologic studies. *Epidemiology*. 2012; 23:350–1. author reply 1-2. [PubMed: 22317815]
37. Suren P, Roth C, Bresnahan M, et al. Association between maternal use of folic acid supplements and risk of autism spectrum disorders in children. *JAMA*. 2013; 309:570–7. [PubMed: 23403681]
38. Haberg SE, Stigum H, Nystad W, Nafstad P. Effects of pre- and postnatal exposure to parental smoking on early childhood respiratory health. *American Journal of Epidemiology*. 2007; 166:679–86. [PubMed: 17625224]
39. Susser M, Stein Z. Timing in prenatal nutrition: a reprise of the Dutch Famine Study. *Nutrition Reviews*. 1994; 52:84–94. [PubMed: 8015751]
40. Lawlor DA, Relton C, Sattar N, Nelson SM. Maternal adiposity—a determinant of perinatal and offspring outcomes? *Nature Reviews Endocrinology*. 2012; 8:679–88.
41. Lawlor DA. The Society for Social Medicine John Pemberton Lecture 2011. Developmental overnutrition—an old hypothesis with new importance? *International Journal of Epidemiology*. 2013; 42:7–29. [PubMed: 23508404]
42. Ebbeling CB, Pawlak DB, Ludwig DS. Childhood obesity: public-health crisis, common sense cure. *Lancet*. 2002; 360:473–82. [PubMed: 12241736]
43. Whitaker RC, Wright JA, Pepe MS, Seidel KD, Dietz WH. Predicting obesity in young adulthood from childhood and parental obesity. *New England Journal of Medicine*. 1997; 337:869–73. [PubMed: 9302300]
44. Lawlor DA, Davey Smith G, O'Callaghan M, et al. Epidemiologic evidence for the fetal overnutrition hypothesis: findings from the mater-university study of pregnancy and its outcomes. *American Journal of Epidemiology*. 2007; 165:418–24. [PubMed: 17158475]
45. Lawlor DA, Timpson NJ, Harbord RM, et al. Exploring the developmental overnutrition hypothesis using parental-offspring associations and FTO as an instrumental variable. *PLoS Medicine*. 2008; 5:e33. [PubMed: 18336062]
46. Davey Smith G, Steer C, Leary S, Ness A. Is there an intrauterine influence on obesity? Evidence from parent child associations in the Avon Longitudinal Study of Parents and Children (ALSPAC). *Archives of Disease in Childhood*. 2007; 92:876–80. [PubMed: 17595200]
47. Kivimaki M, Lawlor DA, Davey Smith G, et al. Substantial intergenerational increases in body mass index are not explained by the fetal overnutrition hypothesis: the Cardiovascular Risk in Young Finns Study. *The American Journal of Clinical Nutrition*. 2007; 86:1509–14. [PubMed: 17991666]
48. Patel R, Martin RM, Kramer MS, et al. Familial associations of adiposity: findings from a cross-sectional study of 12,181 parental-offspring trios from Belarus. *PLoS ONE*. 2011; 6:e14607. [PubMed: 21298034]
49. Fleten C, Nystad W, Stigum H, et al. Parent-offspring body mass index associations in the Norwegian Mother and Child Cohort Study: a family-based approach to studying the role of the intrauterine environment in childhood adiposity. *American Journal of Epidemiology*. 2012; 176:83–92. [PubMed: 22771730]
50. Vik KL, Romundstad P, Carslake D, Davey Smith G, Nilsen TI. Comparison of father-offspring and mother-offspring associations of cardiovascular risk factors: family linkage within the population-based HUNT Study, Norway. *International Journal of Epidemiology*. 2014; 43:760–71. [PubMed: 24366488]
51. Pembrey ME, Bygren LO, Kaati G, et al. Sex-specific, male-line transgenerational responses in humans. *EJHG*. 2006; 14:159–66. [PubMed: 16391557]
52. Northstone K, Golding J, Davey Smith G, Miller LL, Pembrey M. Prepubertal start of father's smoking and increased body fat in his sons: further characterisation of paternal transgenerational responses. *EJHG*. 2014 Apr 2. [Epub ahead of print]. doi: 10.1038/ejhg.2014.31
53. Davey Smith G. Epigenesis for epidemiologists: does evo-devo have implications for population health research and practice? *International Journal of Epidemiology*. 2012; 41:236–47. [PubMed: 22422459]

54. Dabelea D, Hanson RL, Lindsay RS, et al. Intrauterine exposure to diabetes conveys risks for type 2 diabetes and obesity: a study of discordant sibships. *Diabetes*. 2000; 49:2208–11. [PubMed: 11118027]
55. Lawlor DA, Lichtenstein P, Langstrom N. Association of maternal diabetes mellitus in pregnancy with offspring adiposity into early adulthood: sibling study in a prospective cohort of 280,866 men from 248,293 families. *Circulation*. 2011; 123:258–65. [PubMed: 21220735]
56. Lawlor DA, Lichtenstein P, Fraser A, Langstrom N. Does maternal weight gain in pregnancy have long-term effects on offspring adiposity? A sibling study in a prospective cohort of 146,894 men from 136,050 families. *American Journal of Clinical Nutrition*. 2011; 94:142–8. [PubMed: 21562086]
57. Kral JG, Biron S, Simard S, et al. Large maternal weight loss from obesity surgery prevents transmission of obesity to children who were followed for 2 to 18 years. *Pediatrics*. 2006; 118:e1644–9. [PubMed: 17142494]
58. Smith J, Cianflone K, Biron S, et al. Effects of maternal surgical weight loss in mothers on intergenerational transmission of obesity. *Journal of Clinical Endocrinology and Metabolism*. 2009; 94:4275–83. [PubMed: 19820018]
59. Dodd JM, Turnbull DA, McPhee AJ, Wittert G, Crowther CA, Robinson JS. Limiting weight gain in overweight and obese women during pregnancy to improve health outcomes: the LIMIT randomised controlled trial. *BMC Pregnancy and Childbirth*. 2011; 11:79. [PubMed: 22026403]
60. Briley AL, Barr S, Badger S, et al. A complex intervention to improve pregnancy outcome in obese women; the UPBEAT randomised controlled trial. *BMC Pregnancy and Childbirth*. 2014; 14:74. [PubMed: 24533897]
61. Shenkin SD, Starr JM, Deary IJ. Birth weight and cognitive ability in childhood: a systematic review. *Psychological Bulletin*. 2004; 130:989–1013. [PubMed: 15535745]
62. Lawlor DA, Bor W, O'Callaghan MJ, Williams GM, Najman JM. Intrauterine growth and intelligence within sibling pairs: findings from the Mater-University study of pregnancy and its outcomes. *Journal of Epidemiology and Community Health*. 2005; 59:279–82. [PubMed: 15767380]
63. Matte TD, Bresnahan M, Begg MD, Susser E. Influence of variation in birth weight within normal range and within sibships on IQ at age 7 years: cohort study. *BMJ*. 2001; 323:310–4. [PubMed: 11498487]
64. Lawlor DA, Clark H, Davey Smith G, Leon DA. Intrauterine growth and intelligence within sibling pairs: findings from the Aberdeen children of the 1950s cohort. *Pediatrics*. 2006; 117:e894–902. [PubMed: 16651293]
65. Yang S, Lynch J, Susser ES, Lawlor DA. Birth weight and cognitive ability in childhood among siblings and nonsiblings. *Pediatrics*. 2008; 122:e350–8. [PubMed: 18676521]
66. Frisell T, Oberg S, Kuja-Halkola R, Sjolander A. Sibling comparison designs: bias from non-shared confounders and measurement error. *Epidemiology*. 2012; 23:713–20. [PubMed: 22781362]
67. Donovan SJ, Susser E. Commentary: Advent of sibling designs. *International Journal of Epidemiology*. 2011; 40:345–9. [PubMed: 21450688]
68. Keyes KM, Davey Smith G, Susser E. On sibling designs. *Epidemiology*. 2013; 24:473–4. [PubMed: 23549193]
69. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *International Journal of Epidemiology*. 2003; 32:1–22. [PubMed: 12689998]
70. Davey Smith G, Ebrahim S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ*. 2005; 330:1076–9. [PubMed: 15879400]
71. Davey Smith G, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Medicine*. 2007; 4:e352. [PubMed: 18076282]
72. Davey Smith G. Use of genetic markers and gene-diet interactions for interrogating population-level causal influences of diet on health. *Genes & Nutrition*. 2011; 6:27–43. [PubMed: 21437028]

73. Davey Smith G, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *International Journal of Epidemiology*. 2004; 33:30–42. [PubMed: 15075143]
74. Botto LD, Yang Q. 5,10-Methylenetetrahydrofolate reductase gene variants and congenital anomalies: a HuGE review. *American Journal of Epidemiology*. 2000; 151:862–77. [PubMed: 10791559]
75. Czeizel AE, Dudas I. Prevention of the first occurrence of neural-tube defects by periconceptional vitamin supplementation. *New England Journal of Medicine*. 1992; 327:1832–5. [PubMed: 1307234]
76. Scholl TO, Johnson WG. Folic acid: influence on the outcome of pregnancy. *American Journal of Clinical Nutrition*. 2000; 71:1295S–303S. [PubMed: 10799405]
77. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Human Molecular Genetics*. 2014; 23:R89–R98. [PubMed: 25064373]
78. VanderWeele TJ, Tchetgen EJT, Cornelis M, Kraft P. Methodological Challenges in Mendelian Randomization. *Epidemiology*. 2014; 25:427–35. [PubMed: 24681576]
79. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. *Statistical Methods in Medical Research*. 2007; 16:309–30. [PubMed: 17715159]
80. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Statistics in Medicine*. 2008; 27:1133–63. [PubMed: 17886233]
81. Lawlor DA, Davey Smith G, Mitchell R, Ebrahim S. Adult blood pressure and climate conditions in infancy: a test of the hypothesis that dehydration in infancy is associated with higher adult blood pressure. *American Journal of Epidemiology*. 2006; 163:608–14. [PubMed: 16467415]
82. Evans WN, Ringel JS. Can higher cigarette taxes improve birth outcomes? *J Public Econ*. 1999; 72:135–54.
83. Davey Smith G, Sterne JAC, Fraser A, Tynelius P, Lawlor DA, Rasmussen F. The association between BMI and mortality using offspring BMI as an indicator of own BMI: large intergenerational mortality study. *Brit Med J*. 2009; 339
84. Carlslake D, Fraser A, Davey Smith G, et al. Associations of mortality with own height using son's height as an instrumental variable. *Econ Hum Biol*. 2013; 11:351–9. [PubMed: 22560304]
85. Keyes KM, Davey Smith G, Susser E. Commentary: Smoking in pregnancy and offspring health: early insights into family-based and 'negative control' studies? *International Journal of Epidemiology*. 2014 In press.
86. Lipton, P. *Inference to the Best Explanation*. 2nd edition. London: Routledge; 2004.
87. Govarts E, Nieuwenhuijsen M, Schoeters G, et al. Birth Weight and Prenatal Exposure to Polychlorinated Biphenyls (PCBs) and Dichlorodiphenyldichloroethylene (DDE): A Meta-analysis within 12 European Birth Cohorts. *Environmental Health Perspectives*. 2012; 120:162–70. [PubMed: 21997443]
88. Leventakou V, Roumeliotaki T, Martinez D, et al. Fish intake during pregnancy, fetal growth, and gestational length in 19 European birth cohort studies. *American Journal of Clinical Nutrition*. 2014; 99:506–16. [PubMed: 24335057]
89. Taal HR, Pourcain B, Thiering E, et al. Common variants at 12q15 and 12q24 are associated with infant head circumference. *Nat Genet*. 2012; 44:532–+. [PubMed: 22504419]
90. Freathy RM, Mook-Kanamori DO, Sovio U, et al. Variants in ADCY5 and near CCNL1 are associated with fetal growth and birth weight. *Nat Genet*. 2010; 42:430–U73. [PubMed: 20372150]
91. Bradfield JP, Taal HR, Timpson NJ, et al. A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat Genet*. 2012; 44:526–+. [PubMed: 22484627]
92. Vrijheid M, Casas M, Bergstrom A, et al. European Birth Cohorts for Environmental Health Research. *Environmental Health Perspectives*. 2012; 120:29–37. [PubMed: 21878421]
93. Larsen PS, Kamper-Jorgensen M, Adamson A, et al. Pregnancy and birth cohort resources in Europe: a large opportunity for aetiological child health research (vol 27, pg 393, 2013). *Paediatric and Perinatal Epidemiology*. 2013; 27:505. -
94. Batty GD, Alves JG, Correia J, Lawlor DA. Examining life-course influences on chronic disease: the importance of birth cohort studies from low- and middle- income countries. An overview.

- Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas / Sociedade Brasileira de Biofisica [et al]. 2007; 40:1277–86.
95. Richter LM, Victora CG, Hallal PC, et al. Cohort profile: the consortium of health-orientated research in transitioning societies. *International Journal of Epidemiology*. 2012; 41:621–6. [PubMed: 21224276]
 96. Brion MJ, Zeegers M, Jaddoe V, et al. Intrauterine Effects of Maternal Prepregnancy Overweight on Child Cognition and Behavior in 2 Cohorts. *Pediatrics*. 2011; 127:E202–E11. [PubMed: 21187310]
 97. Tyrrell J, Huikari V, Christie JT, et al. Genetic variation in the 15q25 nicotinic acetylcholine receptor gene cluster (CHRNA5-CHRNA3-CHRNA4) interacts with maternal self-reported smoking status during pregnancy to influence birth weight. *Human Molecular Genetics*. 2012; 21:5344–58. [PubMed: 22956269]
 98. Eaves L, Pourcain B, Davey Smith G, York T, Evans D. Resolving the Effects of Maternal and Offspring Genotype on Dyadic Outcomes in Genome Wide Complex Trait Analysis (“M-GCTA”). *Behavior Genetics*. 2014:1–11.
 99. Palmer TM, Lawlor DA, Harbord RM, et al. Using multiple genetic variants as instrumental variables for modifiable risk factors. *Statistical Methods in Medical Research*. 2012; 21:223–42. [PubMed: 21216802]
 100. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. *International Journal of Epidemiology*. 2013; 42:1134–44. [PubMed: 24062299]
 101. Evans DM, Brion MJA, Paternoster L, et al. Mining the Human Phenome Using Allelic Scores That Index Biological Intermediates. *PLoS Genet*. 2013; 9
 102. Relton CL, Davey Smith G. Epigenetic Epidemiology of Common Complex Disease: Prospects for Prediction, Prevention, and Treatment. *PLoS Medicine*. 2010; 7
 103. Relton CL, Davey Smith G. Is epidemiology ready for epigenetics? *International Journal of Epidemiology*. 2012; 41:5–9. [PubMed: 22422447]
 104. Michels KB. The promises and challenges of epigenetic epidemiology. *Experimental Gerontology*. 2010; 45:297–301. [PubMed: 20034553]
 105. Foley DL, Craig JM, Morley R, et al. Prospects for epigenetic epidemiology. *American Journal of Epidemiology*. 2009; 169:389–400. [PubMed: 19139055]
 106. Ng JW, Barrett LM, Wong A, Kuh D, Davey Smith G, Relton CL. The role of longitudinal cohort studies in epigenetic epidemiology: challenges and opportunities. *Genome Biology*. 2012; 13:246. [PubMed: 22747597]
 107. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. *Annual Review of Nutrition*. 2007; 27:363–88.
 108. Waterland RA. Is Epigenetics an Important Link between Early Life Events and Adult Disease? *Horm Res*. 2009; 71:13–6. [PubMed: 19153498]
 109. Reynolds RM, Jacobsen GH, Drake AJ. What is the evidence in humans that DNA methylation changes link events in utero and later life disease? *Clin Endocrinol*. 2013; 78:814–22.
 110. Heijmans BT, Tobi EW, Stein AD, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105:17046–9. [PubMed: 18955703]
 111. Joubert BR, Haberg SE, Nilsen RM, et al. 450K Epigenome-Wide Scan Identifies Differential DNA Methylation in Newborns Related to Maternal Smoking during Pregnancy. *Environmental Health Perspectives*. 2012; 120:1425–31. [PubMed: 22851337]
 112. Dominguez-Salas P, Moore SE, Baker MS, et al. Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat Commun*. 2014; 5
 113. Godfrey KM, Sheppard A, Gluckman PD, et al. Epigenetic Gene Promoter Methylation at Birth Is Associated With Child's Later Adiposity. *Diabetes*. 2011; 60:1528–34. [PubMed: 21471513]
 114. Relton CL, Groom A, St Pourcain B, et al. DNA Methylation Patterns in Cord Blood DNA and Body Size in Childhood. *PLoS ONE*. 2012; 7
 115. Groom A, Potter C, Swan DC, et al. Postnatal Growth and DNA Methylation Are Associated With Differential Gene Expression of the TACSTD2 Gene and Childhood Fat Mass. *Diabetes*. 2012; 61:391–400. [PubMed: 22190649]

116. Relton CL, Davey Smith G. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *International Journal of Epidemiology*. 2012; 41:161–76. [PubMed: 22422451]
117. Kirkbride JB, Susser E, Kundakovic M, Kresovich JK, Davey Smith G, Relton CL. Prenatal nutrition, epigenetics and schizophrenia risk: can we test causal effects? *Epigenomics*. 2012; 4:303–15. [PubMed: 22690666]
118. Joubert BR. 450K Epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy (vol 120, pg 1425, 2012). *Environmental Health Perspectives*. 2012; 120:A455–A.
119. Joubert BR, Haberg SE, Bell DA, et al. Maternal smoking and DNA methylation in newborns: in utero effect or epigenetic inheritance? *Cancer Epidemiology, Biomarkers & Prevention*. 2014; 23:1007–17.
120. Richmond RC, Woodward G, Gaunt TR, et al. Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: Findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). 2014 Unpublished at date of submission.
121. Binder AM, Michels KB. The causal effect of red blood cell folate on genome-wide methylation in cord blood: a Mendelian randomization approach. *BMC Bioinformatics*. 2013; 14:353. [PubMed: 24305512]
122. Ala-Korpela M, Kangas AJ, Soininen P. Quantitative high-throughput metabolomics: a new era in epidemiology and genetics. *Genome Medicine*. 2012; 4:36. [PubMed: 22546737]
123. Gieger C, Geistlinger L, Altmaier E, et al. Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. *PLoS Genet*. 2008; 4:e1000282. [PubMed: 19043545]
124. Kettunen J, Tukiainen T, Sarin AP, et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat Genet*. 2012; 44:269–76. [PubMed: 22286219]
125. Chadeau-Hyam M, Athersuch TJ, Keun HC, et al. Meeting-in-the-middle using metabolic profiling - a strategy for the identification of intermediate biomarkers in cohort studies. *Biomarkers : Biochemical Indicators of Exposure, Response, and Susceptibility to Chemicals*. 2011; 16:83–8.
126. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010; 466:707–13. [PubMed: 20686565]
127. Roth C, Magnus P, Schjolberg S, et al. Folic acid supplements in pregnancy and severe language delay in children. *JAMA*. 2011; 306:1566–73. [PubMed: 21990300]

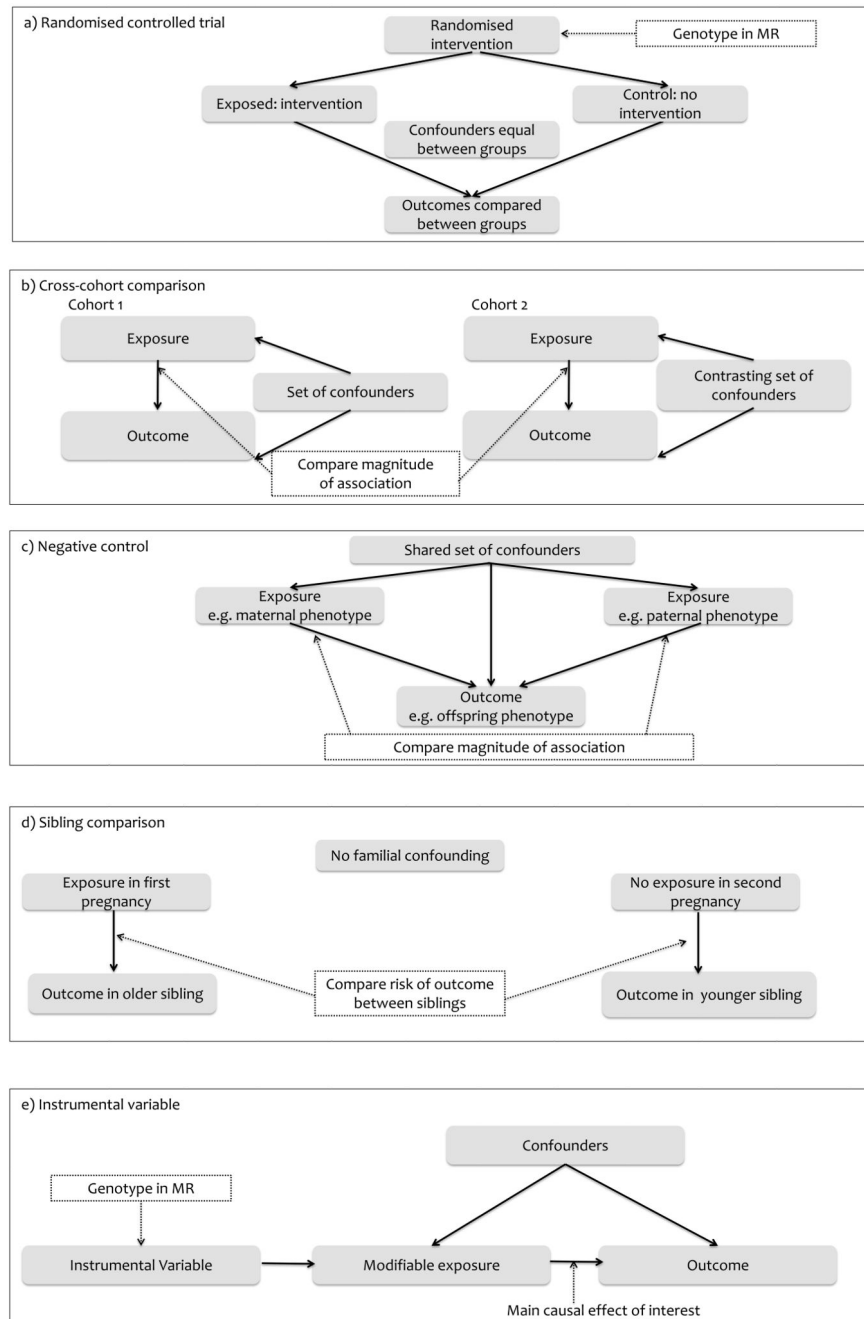


Figure 1.
 Schematic diagrams outlining the main causal inference methods
 MR = Mendelian randomization

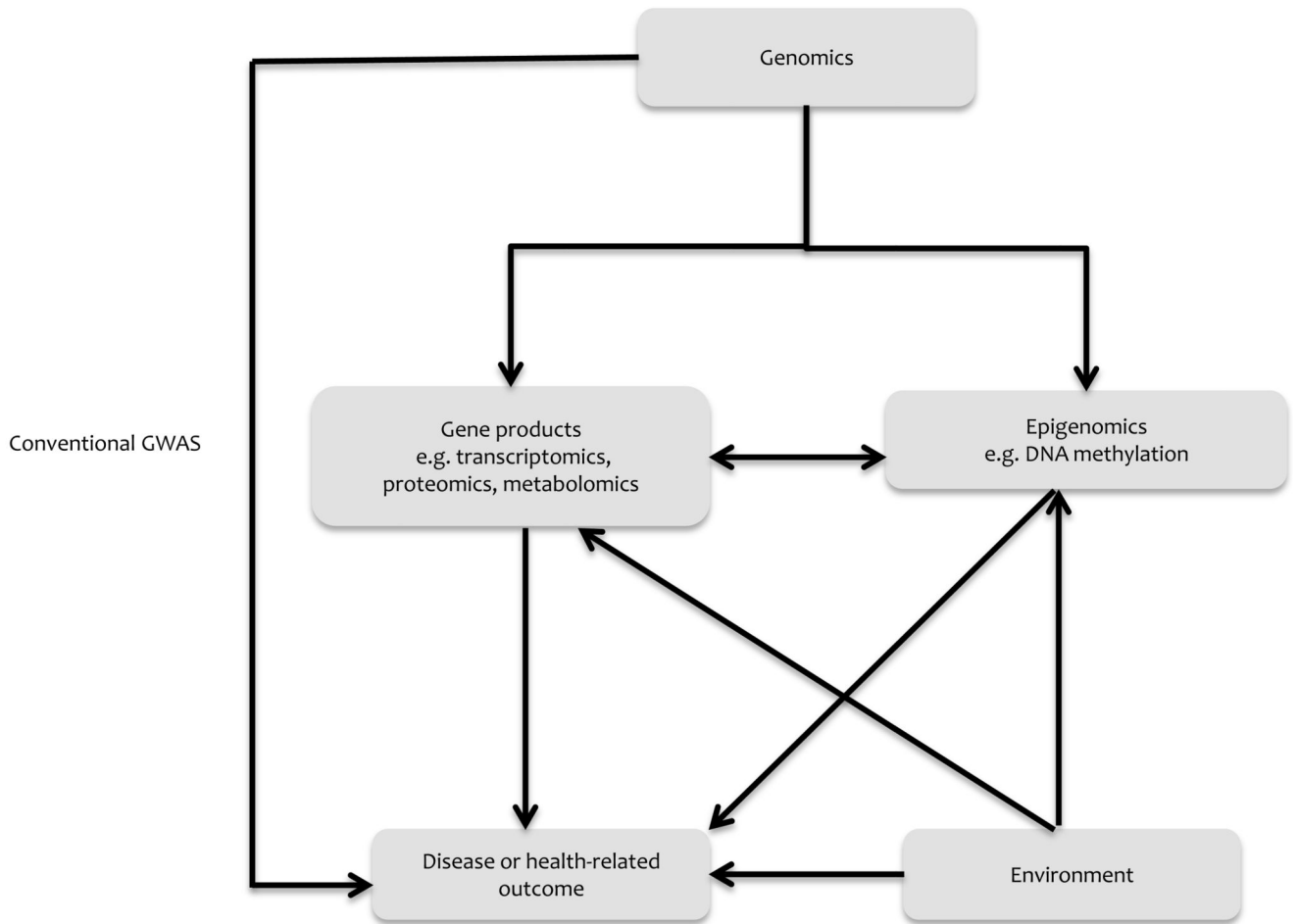


Figure 2. Diagram outlining the interplay between genomics, other “omics” and environmental factors in relation to disease or health-related outcomes
 GWAS = Genome-wide association study

Table 1

METHODS OF DATA COLLECTION	BIOLOGICAL SAMPLES	DNA EXTRACTED	GWAS DATA	EPIGENETIC DATA	METABOLOMIC DATA	SIBLING DATA	DATA ON BOTH PARENTS	PROSPECTIVE/RETROSPECTIVE	FOLLOW-UP	FREQUENCY OF FOLLOW-UP	BIRTH COHORT	STUDY DESCRIPTION	INITIAL SAMPLE SIZE	DATA COLLECTION DURING PREGNANCY
Questionnaires, clinical assessments, medical and educational records	In offspring (at birth and later ages), mothers (during pregnancy and postnatally) and fathers	~10,000 mothers and offspring; ~2,000 fathers	8,365 children, 8,340 mothers, fathers in process	450K DNA methylation available in 7,000 mother-child pairs at multiple time points	NMR metabolomic profiling in 5,000 children, 2,000 mothers and 2,000 fathers planned, at multiple time points	Only if born during recruitment period.	Yes	Prospective	From around 8th gestational week to 21+ years postnatal for parents and between birth and 21+ years of age for child.	68 data collection time points for child, 19 data collection time points for mother.	Avon Longitudinal Study of Parents and Children (ALSPAC) (UK)	Prospective pregnancy cohort study set from 1991 to 1992 in the south west of England and the surrounding areas.	15,247 pregnancies (14,541 in initial recruitment, 706 added a later time)	Yes
Maternal interview at birth; obstetric data from medical records; parental interviews, medical examinations, school attainment and questionnaires in childhood; interviews and questionnaires for offspring (and parents) in adulthood. Biomedical assessment and blood collection at age 44-45. Data from medical records.	In offspring	~7,500 offspring	~3,000 offspring (7,000 metabochip and immunochip)	MeDIP in subsample	No	No	Yes	Prospective though some retrospective data about pregnancy.	Between birth and 55 years of age	Since 1958, 9 further 'sweeps' of all cohort members	1958 British birth cohort (National Child Development Study) (UK)	Longitudinal study of all children born in England, Scotland and Wales in one week in March 1958. Initially a study on perinatal mortality, but later extended to lifelong tracing of participants.	17,416 births	No, some retrospective data from medical records and perinatal survey.
Maternal questionnaire and offspring anthropometric assessment at birth; follow-up interviews, anthropometric assessments and questionnaires; some national record linkage.	In offspring	~4,000 offspring	~3,500 offspring	No	No	Only if born during recruitment period.	Yes, though limited for partners	Mostly prospective data collection. Retrospective data about pregnancy.	Between birth and 30 years	On entire cohort at age 2, 4, 22 and 30 years but more frequent follow-up of subsamples	Pelotas Birth Cohort Studies (1982) (Brazil)	One of three parallel longitudinal studies of all infants being born to mothers living in Pelotas in 1982, 1993 and 2004.	5,914 births	No, though some retrospective data from perinatal survey
Questionnaires and record linkage with Norwegian Patient Registry	In offspring (at birth), mothers (during pregnancy and at birth) and fathers	~70,000-95,000 offspring, mothers and fathers	~3,000 mothers and children, 14,000 mothers and children and 11,000 fathers available in 2015	450K DNA methylation data available in 1,000 mother-child pairs	No	Includes almost 18,000 pairs of siblings	Yes	Prospective	From pregnancy week 17-18 to date, 7 year and further follow-ups are being processed.	3 questionnaires during pregnancy, further questionnaires at 6 months, 18 months and 3 years.	The Norwegian Mother and Child Cohort Study (MoBa) (Norway)	Prospective pregnancy cohort which enrolled women at week 17-18 in pregnancy from all over Norway between 1999 and 2008, with record linkage	>114,500 children born to >95,000 mothers	Yes
Observational assessments, parental and child interview and questionnaires, records of child health care	In mothers (during pregnancy), children (at birth and age 6) and fathers	~6,500 offspring, 8,000 mothers and 5,000 fathers	Available on ~6,000 children. Not yet scheduled for parents.	450K DNA methylation data available in 1,000 cord blood samples	No	Only if born during recruitment period.	Yes	Prospective	From early pregnancy to date. Data collection at age 10 years will be completed in 2015	3 assessments in the prenatal phase, frequent follow-ups from birth to 48 months and data collection at focus visit at age 5-6 years	Generation R (The Netherlands)	Population-based multi-ethnic birth cohort study which enrolled mothers living in a defined area of Rotterdam, Netherlands with a delivery date between April 2002 and January 2006.	9,778 mothers and children	Yes

METHODS OF DATA COLLECTION	BIOLOGICAL SAMPLES	DNA EXTRACTED	GWAS DATA	EPIGENETIC DATA	METABOLOMIC DATA	SIBLING DATA	DATA ON BOTH PARENTS	PROSPECTIVE/RETROSPECTIVE	FOLLOW-UP	FREQUENCY OF FOLLOW-UP	BIRTH COHORT	STUDY DESCRIPTION	INITIAL SAMPLE SIZE	DATA COLLECTION DURING PREGNANCY
centres, obstetric records Computer-assisted telephone interviews; questionnaires; data from Routine Health Registers and National Hospital Discharge Registry.	In offspring (at birth) and mothers (during pregnancy)	~91,000 samples on mothers	GWAS data available on ~8,000 mothers	450K DNA methylation in ~1,300 offspring of GOYA study on available soon	No	Only if born during recruitment period.	Yes	Largely prospective with some retrospective data collection about medical history.	From early pregnancy to date	2 telephone interviews and 1 food questionnaire in pregnancy, further follow-ups at 6 months, 18 months, 7 years and 11 years. The plan is for the study to be lifelong.	The Danish National Birth Cohort (DNBC) (Denmark)	Prospective cohort study which enrolled pregnant women from all over Denmark between 1996 and 2003, with record linkage	92,274 mothers with a total of 100,418 pregnancies	Yes
Telephone interview, questionnaires and clinical assessment	In offspring	~1,000 offspring	No	450K DNA methylation available on ~600 offspring	No	324 siblings of cases formed the control group for the study	Yes	Largely retrospective	One follow up in adulthood	Follow-up took place once between 2003-2005 at around age 59 years.	The Dutch Hunger Winter Families Study (The Netherlands)	Longitudinal cohort study of part of the population that was in the fetal stage at the time of the 'Hunger Winter' of 1944-45. Aimed to examine how maternal under-nutrition during specific gestational time windows may affect the subsequent life course.	3,307 live birth singletons at 3 institutions in famine-exposed cities in Western Netherlands were identified. 1,075 (751 cases and 324 sibling controls) not lost to follow-up.	No, apart from information on stage of pregnancy when famine may have been experienced
Polyclinic and home visits for observational assessment, interviews, parent and child-reported questionnaires and administered tests; data from routine check-ups abstracted from medical records, teacher-reported questionnaires	In offspring at age 11.5	No	No	No	No	No	Yes	Prospective though sometimes retrospective if participant missed one or more follow-up study visit.	From early post-partum to date. Data from age 16 years is currently being processed.	After post-partum survey, 6 further follow-ups took place before 12 months, one at 6.5 years and one at 11.5 years	Promotion of Breastfeeding Intervention Trial (PROBIT) (Belarus)	Cluster-randomised control trial initiated in Belarus to investigate the influence of a breastfeeding promotion intervention. Recruitment from 1996-1997 and mothers eligible if they intended to breastfeed and had given birth to a healthy singleton infant.	17,046 mother-infant pairs	No

Table 2

CAUSAL INFERENCE APPROACH	APPROACH SUMMARY	BIASES ADDRESSED	STRENGTHS	POTENTIAL LIMITATIONS	EXAMPLE	SCHEMATIC
RANDOMIZED CONTROL TRIAL (RCT)	Subjects are randomly allocated to either exposure or control groups with assumption that there is no difference between the two groups except for the intervention they are receiving.	Confounding, reverse causality, selection bias, loss-to-follow up bias (using intention-to-treat analysis), measurement error	Gold standard for estimating causal effects. Any effect is very likely to be causal if study has large number and trial is reliably performed.	Generalizability may be questionable; impossible or unethical to randomize to certain exposures; can be expensive	Association between maternal breastfeeding behaviour and childhood outcomes 27-33	a
CROSS-COHORT COMPARISON	Associations are compared between two or more populations with markedly different confounding structures. If the observed association is causal, it should be present in both cohorts.	Confounding, Selection bias	Exploring residual confounding. Reliable findings if cohorts are markedly different and have large sample size.	Assumptions about different confounding structures may not be correct; need for variable harmonisation between cohorts	Association between breastfeeding and IQ, obesity and blood pressure in two cohorts 34	b
NEGATIVE CONTROL APPROACHES (approach used to rule out possible non-causal interpretation of results by performing study where hypothesised causal mechanism is removed; expected to produce null result)	Effect of an exposure is compared with the effect of another exposure with similar confounding. Causal inference is strengthened if an association is seen with the exposure being tested and not with the similar exposure.	Confounding	Provides information on the specificity of the exposure being tested, and whether the observed association is simply due to confounding by other associated factors and study biases.	Assumption about the structure of confounding for negative control, some uncontrolled confounders	Association between maternal use of folic acid supplements (and comparison with fish oil) and risk of autism in children 37	c
CONTROL EXPOSURE	An outcome that shares similar confounding as the original outcome, but is unlikely to be influenced by the exposure of interest. Causal inference is strengthened if an association is seen with the outcome being tested and not with the similar outcome.	Confounding	Provides information on the specificity of the outcome being tested, and whether the observed association is simply due to confounding by other associated factors and study biases.	Assumption about the structure of confounding for negative control, some uncontrolled confounders	Association between use of folic acid supplements in pregnancy and severe language delay in children (and comparison with motor skills) 127	
CONTROL OUTCOME						

CAUSAL INFERENCE APPROACH	APPROACH SUMMARY	BIASES ADDRESSED	STRENGTHS	POTENTIAL LIMITATIONS	EXAMPLE	SCHEMATIC
NATURAL EXPERIMENT	Empirical study approach where a population is exposed to an external event or intervention at a specific time point. Associations are then compared with a similar cohort who was not exposed. The assumption is that exposure is caused by quasi-random assignment.	Confounding, reverse causality	Can study settings which would be impractical or unethical to produce by researchers	Selection bias as exposure cannot be manipulated by researcher, some unobserved confounding may remain	Association between imposed nutritional deprivation during early pregnancy on a number of health outcomes, compared with those who experienced famine at other stages in pregnancy ³⁹	
PARENTAL COMPARISON	Maternal-child association is compared with paternal-child association for inferring causal effect of intrauterine exposure. If causal, maternal association is stronger than paternal association. Where associations are similar for both parents we assume that they are driven by genetic or postnatal environmental characteristics.	Confounding	Improves causal inference of intrauterine effect if exposures are measured in both parents at same time in pregnancy, and non-paternity is taken into account for phenotypic traits	Assumption that paternal exposures share same confounding structure as maternal exposure may not be correct; where parental associations are similar magnitude this may be due to offsetting paternal pathways rather than shared confounding	Association between parental smoking and birth weight ¹⁸	
SIBLING COMPARISON	Compares outcomes when siblings are discordant for an exposure. If causal then there will be evidence of a difference in outcome in relation to discordant exposure levels within sibships.	Confounding	Improves causal inference of intrauterine exposures. Controls for familial background and related confounding factors.	Assumes a stable family environment; confounding by factors not perfectly shared by siblings; potential for measurement error of exposure; limited power	Effect of maternal gestational diabetes on height, weight and BMI of offspring and their siblings ⁵⁵	d
MENDELIAN RANDOMIZATION (MR)	MR is the use of a genetic variant robustly associated with an exposure/risk factor of interest as an instrumental variable to test and estimate the causal effect of that exposure/risk	Confounding, reverse causality, selection bias, measurement errors, generalizability	Genetic instruments are not subject to confounding from environmental or lifestyle factor; are not	Low power, lack of instrumentation, pleiotropy and linkage disequilibrium, non-population stratification, non-linear associations, developmental canalisation	Association between maternal <i>MTHFR</i> variants and risk of neural tube defects	a, e

CAUSAL INFERENCE APPROACH	APPROACH SUMMARY	BIASES ADDRESSED	STRENGTHS	POTENTIAL LIMITATIONS	EXAMPLE	SCHEMATIC
NON-GENETIC INSTRUMENTAL VARIABLE (IV)	factor with a disease or health related outcome. Similar to MR except that it uses a phenotype rather than a genetic variable as an IV for obtaining and estimating causal effects.	Sometimes confounding, reverse causality	influenced by the outcome, do not change over time and are measured with high accuracy IV is associated with the outcome only through association with exposure and will typically not be associated with confounding factors	True exogenous factors are generally rare or of small effect. When exposure in one family member is used as IV for exposure in another family member, residual confounding is likely.	(NTD) in offspring 74 Association between own BMI and mortality using son's BMI as an IV 83	e