Psychological Review

Causal Inference Methods for Intergenerational Research Using Observational Data

Leonard Frach, Eshim S. Jami, Tom A. McAdams, Frank Dudbridge, and Jean-Baptiste Pingault

Online First Publication, April 24, 2023. https://dx.doi.org/10.1037/rev0000419

AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Review

ISSN: 0033-295X

https://doi.org/10.1037/rev0000419

THEORETICAL NOTE

Causal Inference Methods for Intergenerational Research Using Observational Data

Leonard Frach[1], Eshim S. Jami[1], Tom A. McAdams[2, 3], Frank Dudbridge[4], and Jean-Baptiste Pingault[1, 2]

[1] Department of Clinical, Educational and Health Psychology, Division of Psychology and Language Sciences, Faculty of Brain Sciences, University College London

[2] Social, Genetic and Developmental Psychiatry, Institute of Psychiatry, Psychology & Neuroscience, King's College London

[3] Promenta Research Centre, University of Oslo

[4] Department of Health Sciences, University of Leicester

Identifying early causal factors leading to the development of poor mental health and behavioral outcomes is essential to design efficient preventive interventions. The substantial associations observed between parental risk factors (e.g., maternal stress in pregnancy, parental education, parental psychopathology, parent–child relationship) and child outcomes point toward the importance of parents in shaping child outcomes. However, such associations may also reflect confounding, including genetic transmission—that is, the child inherits genetic risk common to the parental risk factor and the child outcome. This can generate associations in the absence of a causal effect. As randomized trials and experiments are often not feasible or ethical, observational studies can help to infer causality under specific assumptions. This review aims to provide a comprehensive summary of current causal inference methods using observational data in intergenerational settings. We present the rich causal inference toolbox currently available to researchers, including genetically informed and analytical methods, and discuss their application to child mental health and related outcomes. We outline promising research areas and discuss how existing approaches can be combined or extended to probe the causal nature of intergenerational effects.

Keywords: causal inference, genetically informed methods, intergenerational effects, quasi-experimental studies, within-family designs

Understanding the role of parents in shaping their children's outcomes has been a major focus of influential theories across psychology, psychiatry, and behavioral genetics. Delineating the nature of intergenerational effects is however challenging. Randomized controlled trials (RCTs) and experiments are often seen as the gold standard of causal inference but have limited application to intergenerational research due to feasibility and ethical concerns, for example, it would be unethical to design randomized cross-fostering experiments to examine whether environmental or genetic factors explain the intergenerational transmission of risk for schizophrenia. However, observational studies are prone to a variety of biases, including reverse causation, selection biases, or confounding (see Delgado-Rodríguez & Llorca, 2004, for definitions). For example, the outdated theory of the "schizophrenogenic mother" claimed that schizophrenia was caused by an "overprotective but subtly rejecting mother" (Seeman, 2016), which led to social stigma, with maternal behavior being blamed as the cause of their offspring's psychopathology (Harrington, 2012). Research on the "schizophrenogenic" mother was however, distorted by selection and information biases, reverse causation (e.g., using case-only studies and neglecting offspring effects on maternal behavior) and did not account for genetic factors (Parker, 1982). Today, schizophrenia is known to be a

Leonard Frach https://orcid.org/0000-0001-6349-8772
Eshim S. Jami https://orcid.org/0000-0001-7176-2174
Tom A. McAdams https://orcid.org/0000-0002-6825-3499
Frank Dudbridge https://orcid.org/0000-0002-8817-8908
Jean-Baptiste Pingault https://orcid.org/0000-0003-2557-4716

The authors thank Bianca De Stavola for her helpful comments on an earlier version of this article. Leonard Frach, Eshim S. Jami, and Jean-Baptiste Pingault are supported by the European Research Council under the European Union's Horizon 2020 research and innovation program (Grant 863981). Jean-Baptiste Pingault is supported by the Medical Research Foundation 2018 Emerging Leaders 1st Prize in Adolescent Mental Health (MRF-160-0002-ELP-PINGA). Tom A. McAdams is supported by a Wellcome Trust Senior Research Fellowship (Grant 220382/Z/ 20/Z).

No data, materials, or code were used for this article. This review was not preregistered. The authors declare that this work has not been presented or published elsewhere. All authors declare no conflict of interest.

Correspondence concerning this article should be addressed to Leonard Frach, Department of Clinical, Educational and Health Psychology, Division of Psychology and Language Sciences, Faculty of Brain Sciences, University College London, London WC1H 0AP, United Kingdom. Email: leonard.frach.20@ucl.ac.uk

highly heritable disorder with hundreds of common and rare genetic variants identified and studies point toward neurobiological underpinnings linked to synaptic processes and brain maturation starting in fetal life (Birnbaum & Weinberger, 2017; Singh et al., 2022; Trubetskoy et al., 2022). Taken together, those findings suggest that exposure to parental behavior is unlikely to cause development of schizophrenia in the offspring.

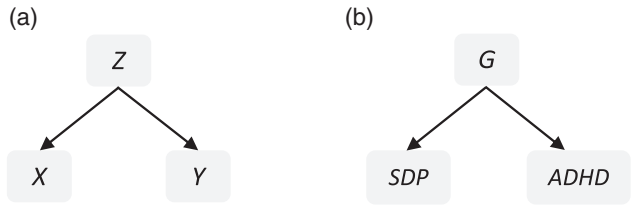
Despite the presence of confounding and other biases in observational research like that on "schizophrenogenic mothers," these biases can be addressed by applying causal inference methods. Resolving the questions of causal effects in turn has fundamental implications not only for preventing adverse and promoting favorable child outcomes, but also for the way we conceive the role of parents in shaping their offspring's development. For example, mistakenly attributing a specific (parental) exposure as the cause of child mental health difficulties would likely result in ineffective interventions due to targeting noncausal factors, thus wasting resources and delaying the development and implementation of more suitable interventions. Furthermore, stigmatizing and blaming either parents or children for being responsible for existing mental health problems could increase mental health problems and negatively impact the parent–child relation.

This review provides a summary of causal inference methods for intergenerational settings that are needed to disentangle putative causal effects from confounded associations, including recent designs using measured genetic variants. For our review, we define exposures as specific parental traits or behaviors with potential effects on a child phenotype. These exposures can either occur prenatally (e.g., maternal stress or drug use during pregnancy), perinatally (e.g., birth complications), or postnatally (e.g., parental education, parent–child relationship quality, parental psychopathology). We draw on empirical examples from the fields of child development, child and adolescent psychology and psychiatry, and allied disciplines, highlighting the multidisciplinary nature of intergenerational research and the broad applicability of such methods beyond those disciplines. We do not offer an extensive background to theories of causal inference and mathematical notation of specific estimands but provide key references. A brief conceptualization of causation is followed by a presentation of existing methods applied to intergenerational settings, where we outline the advantages and limitations of each method and provide references for further readings. Finally, we make suggestions for combining and extending existing approaches to strengthen causal inference.

Causal Inference

Research on the etiology of mental health and behaviors often aims to identify causal effects. An exposure can be considered as a cause for an outcome if changing this exposure results in a change in the outcome of interest, while holding everything else constant. Confounding is a major challenge for causal inference. In the simpler setting of a point-exposure X (i.e., time-fixed exposure), this occurs when a third variable (Z) has a causal effect on both the exposure (X) and the outcome (Y), leading to a spurious association between X and Y (Figure 1a). For example, the effect of smoking during pregnancy on offspring behavioral problems may be entirely confounded by genetic factors predisposing to both smoking behavior and psychopathology (Figure 1b).

Figure 1
Confounding



Note. (a) Visualizlation of confounding, where X has no causal effect on Y. A confounder Z has a causal effect on both X and Y. Not adjusting for Z would result in a spurious association between X and Y. Note that there could also be a real effect ($X \rightarrow Y$), which is only partly confounded by Z. (b) Example of the association between smoking during pregnancy (SDP) and attention deficit hyperactivity disorder (ADHD) which is confounded by shared genetic effects (G), influencing both maternal smoking and child ADHD. For simplification, we assume continuous variables.

Exchangeability

A key assumption of causal inference methods is exchangeability, that is, the same outcome distribution would be observed if exposed and unexposed individuals were exchanged (Greenland & Robins, 1986). Consider the example of smoking during pregnancy as a dichotomous exposure and offspring attention deficit hyperactivity disorder (ADHD) as a dichotomous outcome, where exposure to smoking in utero is assumed to lead to ADHD in childhood. In the counterfactual framework, an individual can be considered as exposed and unexposed at the same time, resulting in two counterfactual outcomes (also termed potential outcomes), for example, developing ADHD or not. The counterfactual outcomes for an individual can be compared to estimate the causal effect of the exposure, for example, a child developing ADHD when exposed to maternal smoking versus not developing ADHD when not exposed (Little & Rubin, 2000).

However, the assumption of exchangeability is not testable in practice, where each individual cannot be exposed and unexposed at the same time. Hence, causal effects cannot be estimated at the individual level due to missing data (Hernán & Robins, 2020). Causal inference methods therefore aim to approximate the counterfactual scenario to meet the exchangeability assumption at the group level. In our example, exchangeability is fulfilled when the distribution of ADHD diagnoses of individuals exposed to smoking during pregnancy would equal the distribution of ADHD in the unexposed individuals if the latter had been exposed to smoking during pregnancy, and vice versa. Concretely, this means that confounders are balanced in the exposed and unexposed groups, so that the treatment is the only meaningful difference. Note that although we present an example with a dichotomous exposure and outcome, the potential outcomes can also be compared for nondichotomous exposures and outcomes (Hernán, 2004).

Exchangeability is achieved in RCTs or randomized experiments by random assignment to the exposed and unexposed groups, resulting in balanced confounders at the group level (Cornfield, 1976). Note that exchangeability may not be present even in RCTs, for example if the randomization was not successful (e.g., small sample size) or because of differential attrition between the groups. Exchangeability

is assumed in other causal inference methods which approximate the counterfactual framework to derive an average causal effect for a given population (Hernán & Robins, 2020). Essentially, designs without random assignment assume exchangeability or conditional exchangeability, that is, individuals are exchangeable after conditioning on a sufficient set of confounders (invoking the strong assumption of no unmeasured confounding). For example, offspring of mothers who smoke during pregnancy and of those who do not may be exchangeable after conditioning on the genetic risk for smoking and ADHD, assuming no other common factors of *X* and *Y* (Figure 1b). There are several ways to account for genetic risk in such examples, which are discussed later in this article. Of note is that in general, different causal effects of an exposure can be estimated, for example, average causal effect for the population or the average causal effect for the exposed. Goetghebeur et al. (2020) highlight the importance of defining the population of interest for which the causal effect should be tested to choose the right method for estimation. This is further discussed elsewhere (Goetghebeur et al., 2020; Hernán, 2004; Hernán & Robins, 2020). Other assumptions for causal inference include positivity, that is, a positive probability for an individual of being exposed and unexposed and consistency, that is, the observed outcome of an individual being exposed equals the counterfactual outcome of being exposed (Hernán & Robins, 2020).

In the current review, we distinguish between causal inference methods that approximate the counterfactual framework and rely on exchangeability by either (a) using random variation induced in the exposure or (b) by implementing statistical or design-based methods that adjust for confounding resulting from nonrandom variation in the exposure (Table 1). Similar distinctions of "instrument-based" and "confounder-control" methods have been used by others (e.g., Matthay et al., 2020), but to distinguish the methods in our review, we introduce a second dimension, that is, natural versus artificial occurrence. "Natural occurrence" in this case means not created for research purposes, for example, a natural disaster, the introduction of a new policy or the twin design. An overview of methods and key limitations are shown in Table 2.

Causal Inference by Using Random Variation in the Exposure

There are numerous quasi-experimental designs which rely on random variation in the exposure to study causal intergenerational effects. As we focus on observational research in this review, we only present such methods that try to achieve exchangeability by naturally occurring means.

Natural Experiments

Naturally occurring events like policy changes or natural disasters can be employed by researchers as quasi-experiments that randomly allocate individuals to given circumstances (Rutter, 2007), which are accompanied by differential exposure to the changing factors (e.g., stress or education). For example, exposed and unexposed individuals can be compared before and after the events. Therefore, comparable to RCTs, natural experiments do not require researchers to identify and control for multiple confounders, although further adjustment for confounders by matching or regression has been recommended (Dunning, 2008).

Natural experiments have also been used to study intergenerational effects. For example, a study examining pregnant mothers and their future offspring following a strong earthquake in Chile reported a negative effect of the prenatal exposure on cognitive outcomes of offspring from disadvantaged families (Torche, 2018). The authors used a difference-in-differences design, which allowed them to compare the difference in outcomes of exposed and unexposed mother–child dyads over time (change/difference over time), suggesting a causal effect of stress induced during the prenatal development of the offspring on cognitive outcomes. Another study used a casino opening in the United States as a natural experiment to study the effect of boosted income in an indigenous population. The study found intergenerational effects, in that a reduction in parental poverty improved some mental health outcomes of their children (Costello et al., 2003).

A major strength of natural experiments is that they allow the study of causal effects of exposures that cannot be administered in an experimental setting. However, the application of natural experiments is limited in that only specific naturally occurring exposures can be studied. Generalizability of findings is also limited (O'Connor, 2003)—for example, the effects of a policy change may be specific for a certain country or certain events might be unique, making similar follow-up research difficult. Not least, the choice of adequate control groups is challenging, as compared groups should closely resemble the exposed group and they should only differ because of the "naturally occurring" exposure. For example, the effect of a specific policy change may be difficult to isolate if other changes occurred around that time.

Regression Discontinuity

Originally introduced as an alternative to RCTs, the regression discontinuity design involves individuals that are assigned to different groups based on a predefined cutoff for an assignment variable (Thistlethwaite & Campbell, 1960), for example based on date of

Table 1
Classification of Causal Inference Methods

	Exchangeability	
Means	(a) Capitalizing on random variation induced in the exposure	(b) Adjusting for confounding resulting from nonrandom variation in the exposure
Naturally occurring	Natural experiments, instrumental variable analysis, Mendelian randomization, regression discontinuity	In vitro fertilization, adoption, twin, sibling, and other family designs
Artificially occurring	Randomized controlled trials, randomized experiments	Doubly robust methods, generalized (g) methods

Note. Broader definitions of natural experiments often include twin studies and other within-family designs (Rutter, 2007), which are also classified as naturally occurring here but in a separate column. In contrast to the other design-based methods, doubly robust and g-methods are analytical methods that statistically control for confounding.

Table 2
Causal Inference Methods, Their Rationale to Meet the Exchangeability Assumption, and Important Limitations

Type	Design	Rationale	Limitations
Random variation in the exposure	Natural experiments	As-if random natural variation in the exposure, for example, following a natural disaster or policy changes. Affected individuals are often compared with one or more control groups. Resembles a random experiment and does not require specific control for confounders.	• limited generalizability of findings • no available data for many exposures and outcomes • choice of adequate control group(s) is a challenge
	Regression discontinuity	Comparing two groups of individuals that are "on the edge" of belonging to the other group, so that compared individuals are very similar to each other but differ in group status (e.g., in treatment status).	• sensitive to misspecification • choice of bandwidth around the cutoff is challenging • treatment assignment based on more than one indicator can be problematic
	Instrumental variable	IV is an unconfounded proxy for the exposure, so that associations of the IV with the outcome are more likely to reflect causal effects of the exposure.	• IV estimation with small sample sizes is imprecise, and with large sample sizes, bias is introduced when assumptions are slightly violated • less useful for cases of strong confounding • adequate instruments may be hard to find for most exposures
	Intergenerational MR	Using genetic instruments, taking advantage of random allocation of genetic variants at birth within families.	• requires large sample sizes and good IVs (strong and robust associations with exposure) • requires genotypes of parents and offspring
Control-based methods	Adoption design	Effects from genetically unrelated adoptive parents on the offspring are free from genetic confounding.	• limited generalizability of findings • adoption at birth necessary to exclude any postnatal nurturing effects of biological parents
	Assisted conception	Comparing effects between genetically related parent–child and genetically unrelated parent–child dyads/trios.	• limited generalizability of findings • small sample sizes
	Parent comparison	Risk factor is specific for one parent and the other one is used as a negative control. This might strengthen causal inference (e.g., intrauterine effects like smoking, BMI, etc.), because other familial confounding is controlled.	• cannot entirely rule out genetic confounding • similarities between parents due to assortative mating or other effects complicate the comparison
	Multiple relationships	Comparing relations in nonnuclear families, including stepparents. Associations between stepparent and child are not genetically confounded, whereas associations between the child and the biological parent with whom the child does not live with mostly reflect genetic effects.	• rearing effects can be underestimated, as time lived with the stepparent can substantially differ from time lived with the biological parent • nurturing effects of not-lived-with biological parents may not be ruled out (e.g., perinatal effects).
	Sibling comparison	Comparing an affected child (e.g., mother smoked during pregnancy) with their unaffected sibling, thus adjusting for familial confounding (shared environment and shared genetics).	• confounders that vary between children are not adjusted for • carryover effects (e.g., birth order) may be present
Control-based methods	Twin comparison	For DZ twins, this design is equivalent to the sibling comparison, but shared effects are better controlled for as twins are the same age. For MZ twins, all genetic effects are accounted for.	• confounders that vary between children, such as nonshared environment are not controlled • direction of causation can only be tested in longitudinal designs • parental exposures that do not vary within families (e.g., parental education) cannot be studied
	Children-of-siblings/ children-of-twins	For MZ twins, this design compares associations in parent-offspring dyad versus uncle-offspring or aunt-offspring dyads (who share the same amount of genes with their niece/nephew as the parent, but not the family environment). Further, avuncular correlations can be compared between MZ and DZ twin families.	• Dyadic parental effects (e.g., divorce) cannot be studied • age differences between siblings and cousins need to be considered for age-dependent outcomes (e.g., cognitive performance)
	G-methods	Generalized methods to study time-varying exposures and to account for time-varying confounding.	• unobserved time-varying confounding cannot be controlled • feedback between time-varying prenatal exposures and outcome examined after birth cannot be accounted
	Doubly robust methods	Methods that include an exposure model (e.g., propensity score) and an outcome model to adjust for confounding. Doubly robust, in that it would be sufficient if one of the two models is correctly specified.	• if both the exposure and the outcome models are mis-specified, results will be biased • when the two models are mis-specified, bias is potentially higher compared to other methods

Note. IV = instrumental variable; MR = mendelian randomization; BMI = body mass index; DZ = dizygotic; MZ = monozygotic. In contrast to the other design-based methods, doubly robust and g-methods are analytical methods that statistically control for confounding.

birth as the cutoff for school eligibility, or based on a level of a biomarker as treatment eligibility. By doing this, a selection bias is created, but one that is perfectly controlled. As individuals just below and above the threshold are quasirandomly allocated to the exposed and unexposed groups due to random variability (e.g., measurement error) in the assignment variable, those individuals are expected to have similar levels regarding all observed and unobserved confounders (Bor et al., 2014). The regression discontinuity design is also commonly used to analyze natural experiments, for example, to test causal effects of altered exposures due to policy changes.

Regression discontinuity designs using predefined cutoffs for kindergarten eligibility showed an effect of being relatively young for their grade on elevated risk of children's ADHD diagnoses (Elder, 2010; Evans et al., 2010). To our knowledge, regression discontinuity designs have not been applied to intergenerational effects on psychological outcomes. However, Ali and Elsayed (2018) used the regression discontinuity design to examine the effects of parental education on child health. The authors compared families including parents that started school just before and after an education reform that resulted in a reduction of compulsory schooling by 1 year. They found little evidence for a causal effect of higher paternal education on offspring's nutritional status and no evidence for an effect on child mortality or other measures (Ali & Elsayed, 2018).

The regression discontinuity design assumes exchangeability or continuity—that is, continuous variation of the assignment variable around the cutoff (Oldenburg et al., 2016). This design allows the study of exposures that are closer to "real life" than by using RCTs, but require larger sample sizes than RCTs (Bloom, 2012). Other limitations, which include the potential bias arising from misspecification of the functional form around the cutoff (e.g., assuming a linear form whereas the "true" functional form might be nonlinear), are addressed in more detail elsewhere together with arguments for the generalizability of the effects (Bloom, 2012).

Instrumental Variables Analysis

Instrumental variable (IV) analyses use an IV associated with an exposure of interest as a proxy to study the exposure's effect on an outcome, and account for both observed and unobserved confounding in the exposure-outcome association (Figure 2a).

For example, educational attainment of parents was shown to influence their offspring's educational attainment using cohorts' finishing school during a period of unrest in France in 1968 as the instrumental variable (Maurin & McNally, 2008). The conflicts between students and universities led to a simplification of the national exams, which subsequently allowed more students to attend higher education. The authors first found increased educational attainment of this generation compared to cohorts' finishing school before and after 1968. Strikingly, these effects were transmitted to the next generation, resulting in higher educational attainment of the students' offspring, and providing evidence for a causal effect (Maurin & McNally, 2008).

Notably, IV analysis requires strong assumptions in addition to exchangeability (exchangeability here would mean that the IV is not associated with any confounders of the exposure-outcome association). The exclusion restriction assumption requires that there should not be any direct association between the IV and the outcome, and that the observed effect should result from the pathway through the exposure (i.e., no other pathways from the exposure to the outcome via other potential mediators). Another assumption is the relevance of the IV to the exposure, which can be tested by examining the magnitude of the association between the IV and the exposure. IV analysis is generally robust against unobserved confounding of the exposure-outcome association, but larger degrees of unobserved confounding reduce the strength of the IV and thus likely violate the assumptions of IV analysis, that is, the relevance assumption (Martens et al., 2006). Inferences from IV analysis can be biased even when assumptions are only slightly violated, or when using small sample sizes (Bound et al., 1995; Martens et al., 2006). Furthermore, it may be challenging to find appropriate instruments for most exposures, as the IV needs to be strongly associated with the exposure and exclusively associated with the outcome through the exposure of interest.

Intergenerational Mendelian Randomization

Mendelian randomization (MR) is similar to IV analysis and make use of measured genetic variants as a proxy for an exposure of interest (Didelez & Sheehan, 2007; Smith & Ebrahim, 2003; see Figure 2b). Individuals inherit one of two copies of each genetic variant randomly from both parents, thus enabling the use of genetic variants with established associations with an exposure as instruments for such exposure. Mendelian randomization has been widely used for causal inference and was recently extended to intergenerational settings (Davies et al., 2019; Zhang et al., 2015).

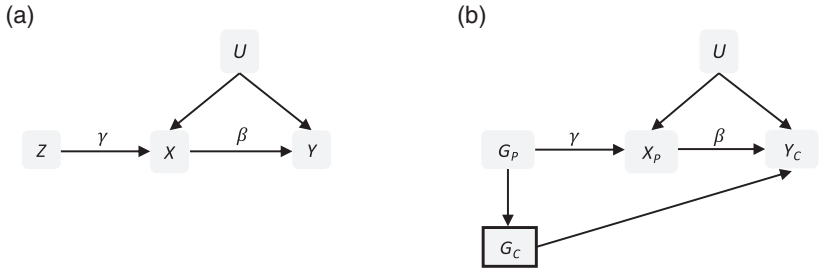
MR was applied using a genetic variant strongly associated with heaviness of smoking as proxy for maternal smoking during pregnancy, to examine the effect of maternal genotype on autism spectrum disorder and related traits of the offspring (Caramaschi et al., 2018). Consistent with other causal inference methods, the MR analysis did not support a causal effect of smoking during pregnancy on the examined outcomes (Caramaschi et al., 2018).

An advantage of MR is that the association between the genetic instrument and the outcome is largely free from environmental confounding but can be influenced by population effects including population stratification or assortative mating. Population stratification describes the fact that certain effect alleles vary in frequency between (sub-)populations, whereas assortative mating indicates that individual mating choices are not random but influenced by phenotypic similarities, for example, tall individuals choosing tall mates (Freedman et al., 2004; Yengo et al., 2018). These biases can be better accounted for in within-family MR analyses (Hwang et al., 2020). For valid inference from intergenerational MR analyses, one needs to include parental genotypes as instruments for the parental exposure and further control for the offspring's genotype (thus requiring data from trios) to block the pathway of genetic transmission (Figure 2b; Davies et al., 2019). Intergenerational MR requires large samples, which are still rare for intergenerational settings. Further limitations include selection bias (Hwang et al., 2020) as well as general limitations of MR and IV analysis such as weak instrument bias (Burgess & Thompson, 2011).

Causal Inference by Using Naturally Occurring Adjustment Methods

As many exposures are not randomly distributed (e.g., whether a child's parents are highly educated or not is usually not random), there is a need for causal inference methods to study potential effects

Figure 2
Instrumental Variable Analysis



Note. (a) General instrumental variable design, where the instrument Z is used as a proxy for the exposure X. Observed associations between the instrument and the outcome (b_{ZY}) are tested to infer the causal effect β, assuming that the observed association b_{ZY} comes exclusively from the pathway $Z \rightarrow X \rightarrow Y$ (exclusion restriction). Conceptually, the exclusion restriction is similar to a mediation analysis, that is, X can be seen as mediator of the association between Z and Y, and this would correspond to a full mediation, where $b_{ZY} = \gamma \times \beta$ and thus the causal effect is $\beta = b_{ZY}/\gamma$. The absence of an arrow between Z and U corresponds to the exchangeability assumption, whereas the assumption of relevance can be tested by the magnitude of γ. (b) Intergenerational Mendelian randomization, using the parental genetic instrument(s) G_P as a proxy for the parental exposure X_P while controlling for the child's genetic variants G_C, as genetic transmission violates the exclusion restriction assumption through the mediating pathway from G_P on Y_C via G_C. C = child; G = genetic instrument; P = parent; U = unmeasured confounder(s); X = exposure; Y = outcome; Z = instrumental variable.

of these exposures using observational data. This can be achieved by confounder control, including methods that control by matching or blocking. Within-family designs are of specific interest when investigating both genetic and environmental effects. Examples of different familial constellations and genetic relatedness in different study designs are shown in Figure 3 and described below.

Adoption Studies

Associations between biological parents' traits or behaviors and the outcomes of their biological offspring might be biased by genetic factors, as parents provide both their genes and the rearing environment for their children. If both the parental exposure and the child outcome are genetically influenced, the association between them can be genetically confounded (e.g., Figure 1b). Adoption designs allow researchers to explore parent–child associations in the absence of potentially confounding effects arising from genetic relatedness. Associations between adoptive parents' traits and the phenotype of the genetically unrelated offspring cannot result from genetic transmission and are thus more likely to reflect environmental transmission (Figure 3a).

A prospective adoption study found no effect of maternal anxiety in the adoptive mother on the anxiety symptoms of the adopted child (Ahmadzadeh et al., 2019), suggesting that previous findings of observational associations between anxiety of biological mothers and anxiety in their biological children (e.g., McClure et al., 2001) may not reflect a causal environmental effect. In contrast, adoption studies support the role of maternal depression as environmental risk factor for offspring internalizing and externalizing problems, as well as for their neurobiological development (reviewed by Natsuaki et al., 2014).

Similar to other designs, the possibility of reverse causation—that is, effects from the child on parental behavior—is a major limitation of this design unless a longitudinal approach is adopted. A specific limitation of adoption studies is that samples may not be representative of the general population, resulting in limited generalizability of findings. Furthermore, although genetic confounding can be ruled out, the design does not control for shared or nonshared environmental confounders (Thapar & Rice, 2020). Previous environmental effects from the biological parents on the offspring's outcomes cannot be ruled out (e.g., during pregnancy or prior to adoption), which also restricts the variety of exposures that can be studied (i.e., only postnatal). However, if additional information on birth mothers and fathers prior to adoption is available (e.g., prenatal environment), maternal effects can also be examined and distinguished from the postnatal rearing environment provided by the adoptive parents. If data on both birth parents are available, prenatal environmental effects can be further distinguished from genetic effects (Loehlin, 2016), which we discuss in the section on parent comparisons.

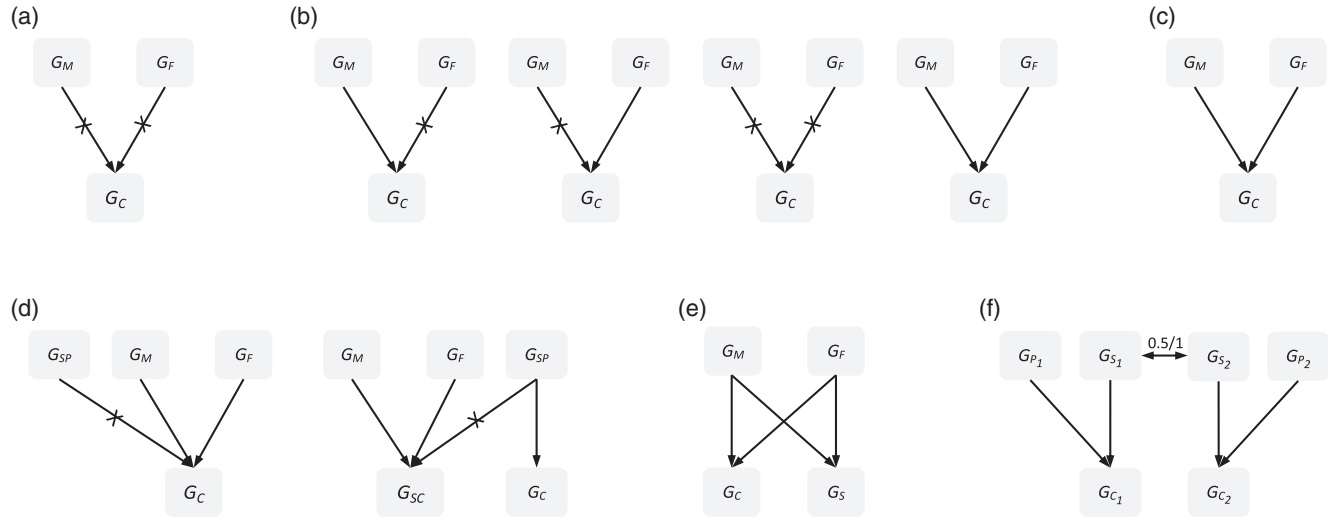
Assisted Conception

Assisted conception or in vitro fertilization (IVF) can be carried out in multiple ways, so that either both parents are biologically unrelated to the child, only one parent is biologically related to the child or both parents are (Figure 3b). Similar to adoption designs, genetic confounding can be ruled out in biologically unrelated parent–child dyads or trios. If the child is carried during pregnancy by the future rearing mother, prenatal effects can be examined in addition to peri- and postnatal effects, which expands the applicability of this design in comparison to the adoption design (e.g., information on the birth mothers may not be available in adoption studies).

Studies using the IVF design showed that maternal smoking during pregnancy is only associated with child ADHD and antisocial behavior in genetically related but not in unrelated families, suggesting genetic confounding of the association between smoking during pregnancy and child psychopathology (Rice et al., 2009; Thapar et al., 2009).

Figure 3

Genetic Relatedness in Different Family Designs



Note. (a) Adoption design; (b) Assisted conception where (i) the sperm is donated (thus the lived-with father is genetically unrelated), (ii) the ovum is donated, (iii) both sperm and ovum are donated, (iv) both sperm and ovum come from the parents seeking IVF; (c) parent comparison; (d) multiple relations; (e) sibling and twin comparisons; (f) children-of-siblings designs, where siblings share either 50% (siblings or DZ twins) or 100% of their genes (MZ twins). Single-headed arrows indicate a directed effect, assuming that a biological parent transmits 50% of their deoxyribonucleic acid to their child, whereas crossed arrows indicate a blocked pathway. Genetic relatedness for the offspring generation can be inferred by path tracing, for example, for Figure 3e, the relatedness between the child and the sibling is equal to the paths via the mother ($G_S \leftarrow G_M \rightarrow G_C$) and father ($G_S \leftarrow G_F \rightarrow G_C$), resulting in a genetic relatedness coefficient of $0.5 \times 0.5 + 0.5 \times 0.5 = 0.5$. Similarly, in Figure 3f the correlation between G_{C1} and G_{C2} is $0.5 \times 1 \times 0.5 = 0.25$ when the parents are MZ twins, similar to half-siblings, and $0.5 \times 0.5 \times 0.5 = 0.125$ when the parents are DZ twins or siblings; similar to typical cousins. C = child; F = father; G = genotype; M = mother; P_1/P_2 = Partner 1/Partner 2; S = sibling; SC = stepchild; SP = stepparent; IVF = in vitro fertilization; DZ = dizygotic; MZ = monozygotic.

Limitations of the assisted conception design are similar to the adoption design, mainly the extent of generalizability, as families using IVF might differ from families conceiving naturally, for example, showing higher rates of perinatal complications or lower degrees of prenatal risk exposure (Thapar & Rice, 2020). However, comparing both biologically related and unrelated families using IVF may control for confounding that might be specific for families using assisted conception. Like adoption studies, present IVF studies are limited by small sample sizes.

Parent Comparison

This design is usually employed for pregnancy exposures, using the father as a negative control to test for a causal intrauterine effect during pregnancy (Figure 3c), but applications to other exposures that are specific to only one parent may be conceivable.

A prospective longitudinal study examined the effect of maternal cannabis use during pregnancy on emotional and behavioral problems of the offspring (El Marroun et al., 2019). They found associations of similar size for maternal and paternal cannabis use during pregnancy with elevated child externalizing problems, providing evidence against a causal intrauterine effect and indicating a role of genetic or other residual confounding (El Marroun et al., 2019).

In the example above, 80% of mothers who used cannabis during pregnancy also had a spouse consuming cannabis, thus maternal cannabis use could not be examined independently (El Marroun et al., 2019). When parental risk factors are highly correlated—reflecting assortative mating or spousal interaction effects—the parental comparison might not be capable of drawing causal inferences. Another limitation of this design is that genetic confounding may not be entirely excluded, as for some disorders sex-specific and sex-dependent genetic effects (Goldstein et al., 2013; Kang et al., 2020) or parent-of-origin effects have been reported. In case of prenatal exposures, results can further be confounded by offspring's larger exposure to maternal versus paternal postnatal environment.

Multiple Relationship Designs

Multiple relationship designs investigate not only nuclear families but focus on families including related individuals not living together, and both related and unrelated individuals living together. Therefore, these designs allow researchers to disentangle genetic and environmental effects and can either be parent-focused—that is, multiple parenting relationships of one (step)parent—or child-focused, that is, multiple relationships between a child and their (step)parents (Figure 3d).

A study examined triparental families comprising children, their biological mother, stepfather and biological father with whom they did not live (Kendler et al., 2015). In investigating parental transmission of different externalizing behaviors, the authors found that maternal effects were largest, followed by effects of not-lived-with biological fathers and then stepfathers, indicating that both environmental and genetic transmission were present (Kendler et al., 2015). Using a similar design, positive effects of higher quality relationships of both stepfather–child and nonresident (biological) father–child relations on child educational attainment were found (King et al., 2020), supporting an environmental effect of parent–child relationships on child education.

These designs are limited—somewhat similar to adoption designs—in that not-lived-with biological parents may in fact live with their offspring in early sensitive stages of their development or often have contact with their child, thus, environmental effects of those parents cannot be ruled out (Kendler et al., 2015, 2019). Furthermore, as stepparents may spend less time with the offspring (e.g., joined the family some years after the child was born), using stepparents as an index for environmental parent-offspring effects can lead to substantial underestimation of rearing effects (Kendler et al., 2015).

Sibling and Twin Comparisons

The sibling comparison contrasts an affected child with their unaffected sibling (i.e., discordant siblings in case of binary exposure), thus controlling for familial confounding including shared genetic and environmental effects (Figure 3e). For instance, siblings may be differentially exposed to risk or protective factors like parental age, parental education, and also parenting behavior. If the outcomes of both siblings are similar, a causal effect of the specific exposure is not likely. A recent article proposed a counterfactual-based framework for the sibling comparison design and describes how a causal effect can be estimated and interpreted (Petersen & Lange, 2020).

Consistent with findings from the IVF design, studies using sibling comparisons suggest that maternal smoking during pregnancy does not causally affect child internalizing and externalizing problems, but that observed associations reflect familial confounding instead (Meier et al., 2017; Obel et al., 2016). Other studies use twin comparisons to examine the associations of differential parent–child relationships with differences in behavioral outcomes of monozygotic twins, which can rule out all genetic and shared environmental confounding (Asbury et al., 2003; Burt et al., 2006). In twin comparison studies, differential parenting behaviors such as harsh discipline have been associated with differences in externalizing and other maladaptive behaviors of the twins (Asbury et al., 2003).

Limitations of these designs are that unobserved nonshared genetic and environmental confounders (i.e., factors that are specific for each child) are not controlled for. Furthermore, the assumption that the direction of effects in parent–child associations is from the parent to the child is not always reasonable (Bell, 1968). Many studies have shown that children can influence their parents (e.g., Ahmadzadeh et al., 2019; Lifford et al., 2008; McAdams et al., 2015). The possibility of reverse causation is therefore something which should be considered, and which can make it more difficult to accurately estimate the effect of the parent on the child. However, this limitation is not unique to the sibling comparisons, but also applies for example, to the adoption design or other designs. Exploiting a longitudinal approach for example, could help to disentangle associations and provide insights in the direction of effects or potential bidirectional effects. Sibling comparisons can also be biased by carryover effects, that is, the exposure of an individual affects the outcome or the exposure of the sibling, but these can be tested (e.g., birth order effects; Sjölander et al., 2016). Furthermore, sibling comparisons cannot be used to study shared parental exposures, for example, household income or parental education, unless such factors vary over time and siblings are differentially affected at a specific age (Sariaslan et al., 2021). In contrast, twin comparisons control for more shared confounding but are even more limited in terms of the exposures that can be studied, for example, precluding the ability to study effects of the prenatal environment or of parental age at birth (D'Onofrio et al., 2014).

Children-of-Siblings Designs

These designs use parents who are related to one another (e.g., twins, siblings, half-siblings) and their children to study associations between parental characteristics and child outcomes, using the cosibling as a matched control to enable causal inference (Figure 3f; Latvala et al., 2015; Rodgers et al., 2008). A special case of this design is the children-of-twins design (D'Onofrio et al., 2003), which we outline here in more detail. Because MZ twins are genetically identical, they are related to their cotwin's child (their niece/nephew) to the same degree as to their own child (sharing 50% of their genes). As a result, children of MZ twins (who are cousins) are biologically half-siblings (they share 25% of their genes on average). In contrast, children of dizygotic (DZ) twins or of siblings share 12.5% of their genes on average. Genetic and environmental effects can be estimated in different ways by comparing associations between twins discordant for a phenotype and their children's phenotype, and associations can be contrasted for children of MZ and DZ twins (D'Onofrio et al., 2003). Using structural equation modeling or other methods, the association between parental and child phenotype can be estimated, while controlling for shared environmental and genetic confounding (McAdams et al., 2018).

The children-of-twins design has been used to examine effects of parental depression on offspring psychopathology. Studies report effects of exposure to parental depression in childhood and adolescence on offspring's depressive symptoms that remain after accounting for shared environment and genetics (McAdams et al., 2015; Silberg et al., 2010; Singh et al., 2011). However, the association between prenatal depression and subsequent child psychopathology, which has been reported in previous observational studies (Barker et al., 2011), does not remain in the children-of-twins design and might be attributable to genetic confounding (Hannigan et al., 2018).

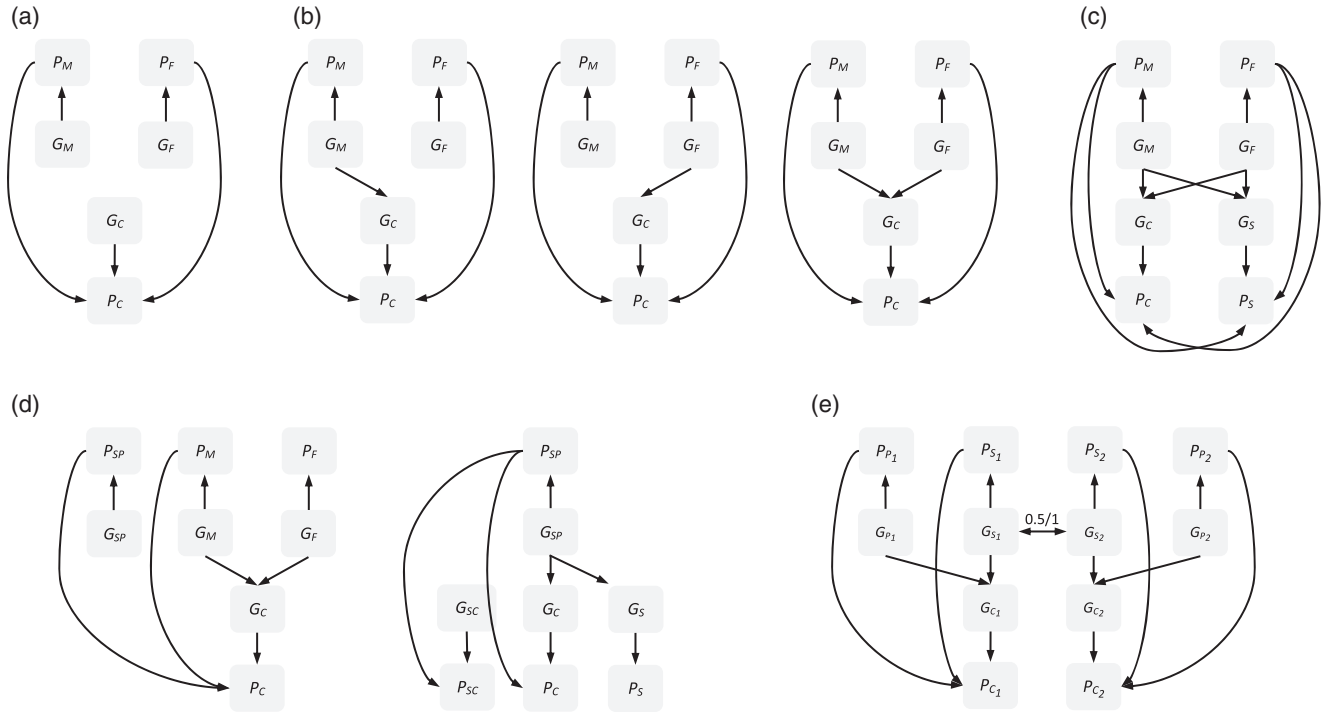
This design does not allow causal inference when studying dyadic effects that cannot be examined for the twins independently but also involve their partners (e.g., divorce or parental conflicts; Eaves et al., 2005). Further limitations and considerations are discussed elsewhere in more detail (D'Onofrio et al., 2003; McAdams et al., 2018).

Genetic Nurture Effects

The concept of genetic nurture describes indirect genetic effects of parental genes on child phenotypes that are assumed to work through environmental pathways, such as parental nurturing (Kong et al., 2018). Polygenic scores—summary scores that capture the genetic predisposition for a trait and include weighted effects from many genetic variants associated with that trait—can be used to infer genetic nurture effects from associations between parental genetic scores and the offspring's outcome (Pingault, Allegrini, et al., 2022). If parental nontransmitted genes are associated with the child phenotype (i.e., ruling out genetic transmission), this effect may occur through the parental environment (Bates et al., 2018; Kong et al., 2018). Genetic nurture effects have been examined recently to investigate associations between parental genotype and offspring educational and mental health outcomes (Jami et al., 2020; Pingault, Barkhuizen, et al., 2022; Wang et al., 2021). Figure 4 is an extension of Figure 3 and additionally includes environmental

Figure 4

Genetic and Environmental Effects in Different Family Designs



Note. (a) Adoption design; (b) assisted conception where (i) the sperm is donated (thus the lived-with father is genetically unrelated), (ii) the ovum is donated, (iii) both sperm and ovum come from the parents seeking IVF (excluding the example where offspring is genetically unrelated to both mother and father as it is identical to Figure 4a); (c) Sibling and twin comparisons; (d) multiple relationships which are either child-focused (left) or parent-focused (right), comparing parent–child relations based on providing either genetics only, environment only or genetics + environment; and (e) children-of-siblings designs. The parent comparison design is not illustrated as it is identical to the third example of Figure 4b. Unlike Figure 3, no (genetic) relation between two individuals is indicated by a missing arrow connecting the two. For simplification, phenotypic effects within a generation (e.g., between siblings or between parents) and effects from child to parent are omitted here. C = child; F = father; G = genotype; M = mother; P = phenotype; P_1/P_2 = Partner 1/Partner 2; S = sibling; SC = stepchild; SP = stepparent; IVF = in vitro fertilization.

parent-to-offspring effects. Furthermore, indirect genetic effects from parental genotype to the offspring's outcome via parental phenotype can be traced in Figure 4 (e.g., $G_M \rightarrow P_M \rightarrow G_C$). Studies designed to test for genetic nurture effects are not prone to reverse causation, as the child cannot influence parental genotypes. However, causal inference from such indirect genetic effects needs to be treated with caution, as one often cannot distinguish between population effects such as population stratification or assortative mating and genetic nurture effects (Young et al., 2019), although new methods can control for some population effects (Balbona et al., 2021; Kim et al., 2021). Furthermore, as genetic nurture effects typically index a global effect without measurement of any specific parental traits (usually only a proportion of genetic variants that are associated with a specific trait are measured), research should be conducted to further characterize genetic nurture effects and to identify environmental mediators. In fact, parental phenotypic data can be integrated in novel approaches to test specific environmental pathways (Balbona et al., 2021; Kim et al., 2021), and a study following up genetic nurture effects on child depression showed partial environmental mediation of these effects by maternal anxiety and depression symptoms (Cheesman et al., 2020). Additionally, specific associations may be followed up using intergenerational MR to examine a causal effect.

Causal Inference Using Control-Based Statistical Methods

In contrast to the above-described methods, which control for unmeasured genetic and familial confounding by study design, many statistical methods aim to control for measured confounders. Among existing methods, some causal inference methods are available that have explicit assumptions for interpreting a causal effect. As mentioned, control-based statistical methods usually assume no unobserved confounding, a strong assumption that can be easily violated. Here we present some exemplar control-based causal inference methods using statistical modeling. Other methods exist such as propensity score methods (i.e., combining measured covariates into one score that captures the propensity of being exposed and adjusting for this score, e.g., by matching or weighting, Austin, 2011), which we did not address here, as they can be more biased or less efficient than g-methods or doubly robust methods (Chatton et al., 2020; Gruber & van der Laan, 2010).

G-Methods

Robins' generalized methods (g-methods) are dedicated to the analysis of longitudinal data with multiple measurements of

(time-varying) exposures. The g-methods include the g-formula, g-estimation of structural nested models, and inverse probability weight estimation of marginal structural models (Robins & Hernán, 2009). Mathematically, these methods use different formulae to estimate the probability densities of outcomes in the exposed and unexposed individuals, given a set of measured confounders (described in detail elsewhere, Robins & Hernán, 2009). In longitudinal designs, individuals can be compared with themselves over time, thus differences between individuals (including differences in confounders) are better accounted for. Other conventional methods, such as fixed effect models, which control for unobserved time-invariant confounding (Gunasekara et al., 2014) are likely to be insufficient to estimate an unbiased causal effect of a time-varying exposure (Robins & Hernán, 2009) as many (parental) exposures and potential confounders are not stable over time, for example, smoking status or income. G-methods can enable estimation of a causal effect of an exposure that varies over time while controlling for time-varying confounding.

For example, a recent study used g-methods (marginal structural models) to examine the effect of maternal smoking during pregnancy on offspring's risk of depression in a sibling design and found similar results compared to generalized estimating equations, reporting an association between maternal smoking and offspring's depression beyond confounding (Shenassa et al., 2023).

Each of the g-methods has specific limitations and one or the other may be favored depending on the data structure and model of interest (Daniel et al., 2013). In general, inverse probability weighting and g-estimation are more robust than the g-formula, although it is possible that all methods converge in some scenarios (Robins & Hernán, 2009). A brief example of g-methods can be found elsewhere (Naimi et al., 2017), including references for more detailed background (e.g., Daniel et al., 2013; Robins & Hernán, 2009). G-methods can also be applied to nonvarying (time-fixed) exposures. However, it may be limited when studying time-varying prenatal exposures on child outcomes that are assessed after birth, as the feedback between exposure and outcome at different stages cannot be examined.

Doubly Robust Methods

Doubly robust methods combine two approaches to control for observed confounding, making them more robust against model misspecification (Funk et al., 2011). Confounders are included in the specification of the outcome-regression model (e.g., as covariates) and of an exposure model, for example, by using propensity scores that describe the probability of being exposed given the observed confounders (Funk et al., 2011). Both the exposure model and the outcome-regression model are prone to misspecification, but combining the two approaches enables a reliable estimation even when one of the two models is mis-specified (Bang & Robins, 2005). A commonly used doubly robust estimator is the targeted maximum likelihood estimator (van der Laan & Rubin, 2006). This method combines maximum likelihood estimation with function-based estimation, where an initial (targeted) density estimator can be updated to maximize the log-likelihood fit (van der Laan & Rubin, 2006).

A study used targeted maximum likelihood estimation to study effects of childhood adversity, providing evidence for an effect of low parental education, neglect and several other forms of adversity on lower fluid intelligence of adolescents (Platt et al., 2018). Recently, a doubly robust method has been introduced to control for confounding by cluster (Zetterqvist et al., 2016), which occurs in family designs due to shared genetics and environment. Applying this method to an existing data confirmed the results from the original study (Kuja-Halkola et al., 2014), providing no evidence of a causal effect of smoking during pregnancy on child cognitive outcomes.

In general, doubly robust methods are limited in that, under correct model specification, they can be less efficient than maximum likelihood estimation and more biased when both models are misspecified (Funk et al., 2011; Kang & Schafer, 2007). Hence expert knowledge and careful model design (including the choice of confounders) is essential.

Sensitivity Analysis

In general, sensitivity analysis is an umbrella term for various post hoc analyses to test the robustness of observed findings. Here we define sensitivity analysis—also termed bias analysis—as a method that specifically tests for potential effects of unmeasured confounding (e.g., Cornfield et al., 1959; Ding & VanderWeele, 2016; Lash et al., 2009). Control-based statistical methods, including the ones described above (g-methods and doubly robust methods), assume no unobserved confounding; however, this assumption cannot be tested and is often unrealistic. The rationale of sensitivity analysis is to test how strong the effect of an unmeasured confounder must be to change the observed finding. Combined with knowledge about effect sizes of existing confounders, one can assess the robustness of the observed association and could draw conclusions about the causal nature of the association.

A study detected an effect of higher prenatal exposure to airborne particles with a diameter <2.5 μm on worsened children's mental and psychomotor development (Tozzi et al., 2019). A sensitivity analysis showed that findings were robust, as effects of unmeasured confounders had to be unrealistically large to erase the observed effect (Tozzi et al., 2019).

Sensitivity analyses may be limited in that they often have strong and untestable assumptions about the frequency or the distribution of the unobserved confounders, albeit some more flexible approaches exist (Ding & VanderWeele, 2016; Shen et al., 2011). Sensitivity analysis still constitutes a valuable tool and complements the above-described methods to strengthen causal inference.

Discussion

Confounding remains a major issue in observational research, including genetic confounding which might partly or fully explain observational associations in intergenerational research, highlighting the importance of causal inference methods. We have described a broad range of methods for causal inference in intergenerational settings, which aim to achieve exchangeability by either (a) capitalizing on random variation in the exposure or (b) adjusting for confounding resulting from nonrandom variation in the exposure. We elaborated on the advantages and limitations of each design and provided an overview of existing methods with references to potential designs for specific research questions.

As each method described here has its own limitations, we advocate combining more than one approach to infer causality. Where feasible, different approaches should be combined in the same study (Caramaschi et al., 2018; Liu et al., 2021), for example, using designs with random variation in the exposure and control-based methods,

performing sensitivity analyses and including negative controls (e.g., pseudo-outcomes; Imbens, 2015). Triangulating evidence from independent studies, that is, using different methods with distinct assumptions to test the same hypothesis, is pivotal to infer causal effects (Lawlor et al., 2016; Munafò et al., 2021; Munafò & Smith, 2018), and no method should be seen as sufficient on its own. Coming back to an example from the introduction (Figure 1b) highlights how a variety of different methods jointly provide insights into the causal relations between intergenerational risk factors and psychopathology. For example, within-family designs using IVF (Thapar et al., 2009), parent (Gustavson et al., 2017; Langley et al., 2012), or sibling comparisons (Knopik et al., 2016; Kuja-Halkola et al., 2014; Skoglund et al., 2014), as well as studies using genetic instruments (Haan et al., 2021) showed that smoking during pregnancy seems to not causally affect child ADHD, but rather reflects familial confounding. Triangulation can also be obtained via systematic reviews and meta-analyses of various genetically informed methods with different assumptions, which have provided evidence for environmental effects of parental anxiety on child internalizing problems (Ahmadzadeh et al., 2021), of prenatal alcohol exposure on child cognitive outcomes (Mamluk et al., 2020), and of several parental risk factors on child mental health and related outcomes (Jami et al., 2021) beyond genetic confounding.

Outlook

The multidisciplinary nature of intergenerational research holds promise for the development of novel methods to further strengthen causal inference. In general, cooperative research efforts are crucial to maximize insights by increasing sample sizes, and by gathering collective expertise across disciplines. However, smaller but more nuanced studies, which include detailed and harmonized phenotyping and longitudinally examine the samples, are also of importance and constitute a large part of the research cited here. In this review, we only focused on studies and methods that test direct effects of parental exposures on child outcomes and did not address mediating or moderating factors, which is a limitation. Identifying mediating pathways from parental (risk) factors to child outcomes or moderating factors of parental effects is clearly important. For example, mediating factors can be examined using causal mediation analysis (Daniel et al., 2015; Imai et al., 2010), and moderators could be tested using a recently proposed framework for causal moderator analysis based on potential outcomes (Dong et al., 2022). In the following sections, we outline some future directions and promising, potential methodological advances for intergenerational causal inference.

New Data for Intergenerational Research

Valuable data for intergenerational research are increasingly evolving, with multigenerational data on a national scale becoming available in terms of birth registries (Van Der Wel et al., 2019), electronic health records (Friedman et al., 2013), or data on education or criminal records. Combining these data provides a tremendous opportunity to gather information about parental characteristics and child outcomes. Furthermore, new data (e.g., through genotyping) are emerging and could be matched with other registries, providing an important basis for the application of MR or the estimation of genetic nurture effects. Other collaborative efforts using large within-family cohorts will result in publicly available summary statistics that can be used for such purposes (Howe et al., 2022). Combining different registry data implies that comprehensive data in a "natural setting" can be used, which is precisely timed and documented, while being representative of the target population. Hence, these data can further be linked to certain events or time periods, such as policy changes, natural disasters such as flooding or pandemics such as COVID-19. However, such detailed data on individuals raise important questions of data privacy, which have to be carefully addressed (e.g., Behrendt et al., 2018).

Spillover From Related Fields

We have seen that certain causal inference methods are underrepresented in intergenerational psychiatry and psychology, such as regression discontinuity or instrumental variable designs, although both are strong designs for causal inference (Kim & Steiner, 2016). This may result from a cautious attitude of researchers in psychology and psychiatry toward *causal* inference (Grosz et al., 2020; Rutter, 2007), potentially coupled with a lack of statistical training in such methods. It has been argued that causal relations are often implicitly assumed by researchers but not explicitly expressed; however, the expression of such assumptions is essential to choose adequate methods and to address their specific biases (Goetghebeur et al., 2020; Grosz et al., 2020; Pingault, Richmond, & Smith, 2022). As mentioned, applying these designs can also be challenging as appropriate instruments might not be available for certain research questions.

Regression discontinuity designs following policy changes were used to examine causal effects of education or alcohol consumption on mental health outcomes (Courtin et al., 2019; Ertan Yörük & Yörük, 2012). Future studies could expand the analysis to the offspring generation of those individuals to study causal effects of parental characteristics on child psychological outcomes as in other designs (Costello et al., 2003; Maurin & McNally, 2008).

Blurring Boundaries Between Classifications

Although in this review we described methods based on distinct categories (Table 1), such categories are not exclusive. First, methods capitalizing on random variation in the exposure (Column 1 of Table 1) may further statistically adjust for confounders (Row 2, Column 2). This has been suggested, for example, for natural experiments (Dunning, 2008) and is common in research using instrumental variables (Vansteelandt & Didelez, 2018). For example, multivariate MR can also be seen as a method to control explicitly for suspected pleiotropic pathways (Burgess & Thompson, 2015). Second, methods using naturally occurring settings to control for confounding (Row 1) may further adjust for confounders to approach conditional exchangeability. This can be done, for example to control for measured nonshared environmental confounders in twin comparison studies, in addition to the adjustment permitted by the natural design (Row 1, Column 2). Furthermore, researchers may be involved more tightly in the planning phase of policy changes (Column 1) to assess data on potential confounders of individuals likely to be affected by the "natural experiment" in a pre–post design.

Advances in Genetically Informed Designs

There is growing interest in and potential of using genetic data in observational studies to strengthen causal inference (Pingault et al., 2018). For instance, a sensitivity analysis has been introduced which

controls for genetic confounding in observational studies with available genotype information of the participants (Pingault et al., 2021). Causal inference methods like the regression discontinuity design could be used to study causal effects of selective prevention programs for children, using parental risk factors as assignment variable (e.g., indexed by genetic instruments). Comparing individuals whose parents score just above or below the threshold (say, "mean genetic risk") would allow us to infer a causal effect of the preventive intervention. Unlike in an RCT, individuals of higher risk would thus all be eligible for the program, instead of using a random assignment.

Alternatively, instead of contrasting individuals with similar values on the assignment variable, recall-by-genotype designs compare individuals at the extreme ends of low versus high genetic risk for a disease. Similar to MR, these designs exploit the random allocation of genetic variants at conception for causal inference (Corbin et al., 2018). Unlike MR, recall-by-genotype studies only focus on a smaller subsample, which can be examined in detail to understand the causal pathway to disease. A translation to the intergenerational setting could be to screen parents regarding their genetic risk toward a particular disorder, and to follow up on their children's outcomes. Thus, one could examine the effects of parental risk factors on the offspring's outcomes in a smaller subpopulation, allowing deep phenotyping of parents and offspring. Similar to intergenerational MR, adjusting for offspring genotype would be necessary to exclude genetic transmission.

There is a large variety of possible methodological extensions using genetic data to study causal environmental effects. However, as genetics represent sensitive individual data, recall-by-genotype, and other designs using genetic screening need careful study design, high transparency, and good communication of the study aims, the meaning of genetic risk, and consequences for the individuals (Beskow et al., 2010; Corbin et al., 2018), especially when children are involved.

Conclusion

The rich toolbox of causal inference methods for observational studies can also be leveraged in intergenerational settings and so far, has successfully contributed to distinctions between putative causal risk factors and confounded associations. Researchers should consider the numerous available designs and methods when conceiving studies to investigate parental factors in observational research and exploit methods from related fields beyond their current repertoire. Besides the broad range of existing methods, promising new approaches to strengthen causal inference are to be expected. Eventually, observational research on intergenerational effects may inform the design of appropriate preventive interventions. Converging evidence suggesting that observed effects of putative parental risk factors are explained by confounding will have practical implications, for example, redesigning interventions to target factors other than those previously considered.

References

Ahmadzadeh, Y. I., Eley, T. C., Leve, L. D., Shaw, D. S., Natsuaki, M. N., Reiss, D., Neiderhiser, J. M., & McAdams, T. A. (2019). Anxiety in the family: A genetically informed analysis of transactional associations between mother, father and child anxiety symptoms. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *60*(12), 1269–1277. https://doi.org/10.1111/jcpp.13068

Ahmadzadeh, Y. I., Schoeler, T., Han, M., Pingault, J.-B., Creswell, C., & McAdams, T. A. (2021). Systematic review and meta-analysis of genetically informed research: Associations between parent anxiety and offspring internalizing problems. *Journal of the American Academy of Child & Adolescent Psychiatry*, *60*(7), 823–840. https://doi.org/10.1016/j.jaac.2020.12.037

Ali, F. R. M., & Elsayed, M. A. A. (2018). The effect of parental education on child health: Quasi-experimental evidence from a reduction in the length of primary schooling in Egypt. *Health Economics*, *27*(4), 649–662. https://doi.org/10.1002/hec.3622

Asbury, K., Dunn, J. F., Pike, A., & Plomin, R. (2003). Nonshared environmental influences on individual differences in early behavioral development: A monozygotic twin differences study. *Child Development*, *74*(3), 933–943. https://doi.org/10.1111/1467-8624.00577

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, *46*(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Balbona, J. V., Kim, Y., & Keller, M. C. (2021). Estimation of parental effects using polygenic scores. *Behavior Genetics*, *51*(3), 264–278. https://doi.org/10.1007/s10519-020-10032-w

Bang, H., & Robins, J. M. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics*, *61*(4), 962–973. https://doi.org/10.1111/j.1541-0420.2005.00377.x

Barker, E. D., Jaffee, S. R., Uher, R., & Maughan, B. (2011). The contribution of prenatal and postnatal maternal anxiety and depression to child maladjustment. *Depression and Anxiety*, *28*(8), 696–702. https://doi.org/10.1002/da.20856

Bates, T. C., Maher, B. S., Medland, S. E., McAloney, K., Wright, M. J., Hansell, N. K., Kendler, K. S., Martin, N. G., & Gillespie, N. A. (2018). The nature of nurture: Using a virtual-parent design to test parenting effects on children's educational attainment in genotyped families. *Twin Research and Human Genetics*, *21*(2), 73–83. https://doi.org/10.1017/thg.2018.11

Behrendt, C.-A., Joassart Ir, A., Debus, E. S., & Kolh, P. (2018). The challenge of data privacy compliant registry based research. *European Journal of Vascular and Endovascular Surgery*, *55*(5), 601–602. https://doi.org/10.1016/j.ejvs.2018.02.018

Bell, R. Q. (1968). A reinterpretation of the direction of effects in studies of socialization. *Psychological Review*, *75*(2), 81–95. https://doi.org/10.1037/h0025583

Beskow, L. M., Linney, K. N., Radtke, R. A., Heinzen, E. L., & Goldstein, D. B. (2010). Ethical challenges in genotype-driven research recruitment. *Genome Research*, *20*(6), 705–709. https://doi.org/10.1101/gr.104455.109

Birnbaum, R., & Weinberger, D. R. (2017). Genetic insights into the neurodevelopmental origins of schizophrenia. *Nature Reviews Neuroscience*, *18*(12), 727–740. https://doi.org/10.1038/nrn.2017.125

Bloom, H. S. (2012). Modern regression discontinuity analysis. *Journal of Research on Educational Effectiveness*, *5*(1), 43–82. https://doi.org/10.1080/19345747.2011.578707

Bor, J., Moscoe, E., Mutevedzi, P., Newell, M.-L., & Bärnighausen, T. (2014). Regression discontinuity designs in epidemiology: Causal inference without randomized trials. *Epidemiology*, *25*(5), 729–737. https://doi.org/10.1097/EDE.0000000000000138

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, *90*(430), 443–450. https://doi.org/10.1080/01621459.1995.10476536

Burgess, S., Thompson, S. G., & the CRP CHD Genetics Collaboration. (2011). Avoiding bias from weak instruments in Mendelian randomization studies. *International Journal of Epidemiology*, *40*(3), 755–764. https://doi.org/10.1093/ije/dyr036

Burgess, S., & Thompson, S. G. (2015). Multivariable Mendelian randomization: The use of pleiotropic genetic variants to estimate causal effects.

American Journal of Epidemiology, *181*(4), 251–260. https://doi.org/10.1093/aje/kwu283

Burt, S. A., McGue, M., Iacono, W. G., & Krueger, R. F. (2006). Differential parent–child relationships and adolescent externalizing symptoms: Cross-lagged analyses within a monozygotic twin differences design. *Developmental Psychology*, *42*(6), 1289–1298. https://doi.org/10.1037/0012-1649.42.6.1289

Caramaschi, D., Taylor, A. E., Richmond, R. C., Havdahl, K. A., Golding, J., Relton, C. L., Munafò, M. R., Davey Smith, G., & Rai, D. (2018). Maternal smoking during pregnancy and autism: Using causal inference methods in a birth cohort study. *Translational Psychiatry*, *8*(1), Article 262. https://doi.org/10.1038/s41398-018-0313-5

Chatton, A., Le Borgne, F., Leyrat, C., Gillaizeau, F., Rousseau, C., Barbin, L., Laplaud, D., Léger, M., Giraudeau, B., & Foucher, Y. (2020). G-computation, propensity score-based methods, and targeted maximum likelihood estimator for causal inference with different covariates sets: A comparative simulation study. *Scientific Reports*, *10*(1), Article 9219. https://doi.org/10.1038/s41598-020-65917-x

Cheesman, R., Eilertsen, E. M., Ahmadzadeh, Y. I., Gjerde, L. C., Hannigan, L. J., Havdahl, A., Young, A. I., Eley, T. C., Njølstad, P. R., Magnus, P., Andreassen, O. A., Ystrom, E., & McAdams, T. A. (2020). How important are parents in the development of child anxiety and depression? A genomic analysis of parent-offspring trios in the Norwegian Mother Father and Child Cohort Study (MoBa). *BMC Medicine*, *18*(1), Article 284. https://doi.org/10.1186/s12916-020-01760-1

Corbin, L. J., Tan, V. Y., Hughes, D. A., Wade, K. H., Paul, D. S., Tansey, K. E., Butcher, F., Dudbridge, F., Howson, J. M., Jallow, M. W., John, C., Kingston, N., Lindgren, C. M., O'Donavan, M., O'Rahilly, S., Owen, M. J., Palmer, C. N. A., Pearson, E. R., Scott, R. A., … Timpson, N. J. (2018). Formalising recall by genotype as an efficient approach to detailed phenotyping and causal inference. *Nature Communications*, *9*(1), Article 711. https://doi.org/10.1038/s41467-018-03109-y

Cornfield, J. (1976). Recent methodological contributions to clinical trials. *American Journal of Epidemiology*, *104*(4), 408–421. https://doi.org/10.1093/oxfordjournals.aje.a112313

Cornfield, J., Haenszel, W., Hammond, E. C., Lilienfeld, A. M., Shimkin, M. B., & Wynder, E. L. (1959). Smoking and lung cancer: Recent evidence and a discussion of some questions. *Journal of the National Cancer Institute*, *22*(1), 173–203. https://doi.org/10.1093/jnci/22.1.173

Costello, E. J., Compton, S. N., Keeler, G., & Angold, A. (2003). Relationships between poverty and psychopathology: A natural experiment. *JAMA*, *290*(15), 2023–2029. https://doi.org/10.1001/jama.290.15.2023

Courtin, E., Nafilyan, V., Glymour, M., Goldberg, M., Berr, C., Berkman, L. F., Zins, M., & Avendano, M. (2019). Long-term effects of compulsory schooling on physical, mental and cognitive ageing: A natural experiment. *Journal of Epidemiology and Community Health*, *73*(4), 370–376. https://doi.org/10.1136/jech-2018-211746

D'Onofrio, B. M., Rickert, M. E., Frans, E., Kuja-Halkola, R., Almqvist, C., Sjölander, A., Larsson, H., & Lichtenstein, P. (2014). Paternal age at childbearing and offspring psychiatric and academic morbidity. *JAMA Psychiatry*, *71*(4), 432–438. https://doi.org/10.1001/jamapsychiatry.2013.4525

D'Onofrio, B. M., Turkheimer, E. N., Eaves, L. J., Corey, L. A., Berg, K., Solaas, M. H., & Emery, R. E. (2003). The role of the children of twins design in elucidating causal relations between parent characteristics and child outcomes. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *44*(8), 1130–1144. https://doi.org/10.1111/1469-7610.00196

Daniel, R. M., Cousens, S. N., De Stavola, B. L., Kenward, M. G., & Sterne, J. A. (2013). Methods for dealing with time-dependent confounding. *Statistics in Medicine*, *32*(9), 1584–1618. https://doi.org/10.1002/sim.5686

Daniel, R. M., De Stavola, B. L., Cousens, S. N., & Vansteelandt, S. (2015). Causal mediation analysis with multiple mediators. *Biometrics*, *71*(1), 1–14. https://doi.org/10.1111/biom.12248

Davies, N. M., Howe, L. J., Brumpton, B., Havdahl, A., Evans, D. M., & Davey Smith, G. (2019). Within family Mendelian randomization studies. *Human Molecular Genetics*, *28*(R2), R170–R179. https://doi.org/10.1093/hmg/ddz204

Delgado-Rodríguez, M., & Llorca, J. (2004). Bias. *Journal of Epidemiology and Community Health*, *58*(8), 635–641. https://doi.org/10.1136/jech.2003.008466

Didelez, V., & Sheehan, N. (2007). Mendelian randomization as an instrumental variable approach to causal inference. *Statistical Methods in Medical Research*, *16*(4), 309–330. https://doi.org/10.1177/0962280206077743

Ding, P., & VanderWeele, T. J. (2016). Sensitivity analysis without assumptions. *Epidemiology*, *27*(3), 368–377. https://doi.org/10.1097/EDE.0000000000000457

Dong, N., Kelcey, B., & Spybrook, J. (2022). Identifying and estimating causal moderation for treated and targeted subgroups. *Multivariate Behavioral Research*. Advance online publication. https://doi.org/10.1080/00273171.2022.2046997

Dunning, T. (2008). Improving causal inference: Strengths and limitations of natural experiments. *Political Research Quarterly*, *61*(2), 282–293. https://doi.org/10.1177/1065912907306470

Eaves, L. J., Silberg, J. L., & Maes, H. H. (2005). Revisiting the children of twins: Can they be used to resolve the environmental effects of dyadic parental treatment on child behavior? *Twin Research and Human Genetics*, *8*(4), 283–290. https://doi.org/10.1375/twin.8.4.283

El Marroun, H., Bolhuis, K., Franken, I. H. A., Jaddoe, V. W. V., Hillegers, M. H., Lahey, B. B., & Tiemeier, H. (2019). Preconception and prenatal cannabis use and the risk of behavioural and emotional problems in the offspring; a multi-informant prospective longitudinal study. *International Journal of Epidemiology*, *48*(1), 287–296. https://doi.org/10.1093/ije/dyy186

Elder, T. E. (2010). The importance of relative standards in ADHD diagnoses: Evidence based on exact birth dates. *Journal of Health Economics*, *29*(5), 641–656. https://doi.org/10.1016/j.jhealeco.2010.06.003

Ertan Yörük, C., & Yörük, B. K. (2012). The impact of drinking on psychological well-being: Evidence from minimum drinking age laws in the United States. *Social Science & Medicine*, *75*(10), 1844–1854. https://doi.org/10.1016/j.socscimed.2012.07.020

Evans, W. N., Morrill, M. S., & Parente, S. T. (2010). Measuring inappropriate medical diagnosis and treatment in survey data: The case of ADHD among school-age children. *Journal of Health Economics*, *29*(5), 657–673. https://doi.org/10.1016/j.jhealeco.2010.07.005

Freedman, M. L., Reich, D., Penney, K. L., McDonald, G. J., Mignault, A. A., Patterson, N., Gabriel, S. B., Topol, E. J., Smoller, J. W., Pato, C. N., Pato, M. T., Petryshen, T. L., Kolonel, L. N., Lander, E. S., Sklar, P., Henderson, B., Hirschhorn, J. N., & Altshuler, D. (2004). Assessing the impact of population stratification on genetic association studies. *Nature Genetics*, *36*(4), 388–393. https://doi.org/10.1038/ng1333

Friedman, D. J., Parrish, R. G., & Ross, D. A. (2013). Electronic health records and US public health: Current realities and future promise. *American Journal of Public Health*, *103*(9), 1560–1567. https://doi.org/10.2105/AJPH.2013.301220

Funk, M. J., Westreich, D., Wiesen, C., Stürmer, T., Brookhart, M. A., & Davidian, M. (2011). Doubly robust estimation of causal effects. *American Journal of Epidemiology*, *173*(7), 761–767. https://doi.org/10.1093/aje/kwq439

Goetghebeur, E., le Cessie, S., De Stavola, B., Moodie, E. E., Waernbaum, I., & the "on behalf of" the topic group Causal Inference (TG7) of the STRATOS initiative. (2020). Formulating causal questions and principled statistical answers. *Statistics in Medicine*, *39*(30), 4922–4948. https://doi.org/10.1002/sim.8741

Goldstein, J. M., Cherkerzian, S., Tsuang, M. T., & Petryshen, T. L. (2013). Sex differences in the genetic risk for schizophrenia: History of the evidence for sex-specific and sex-dependent effects. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, *162*(7), 698–710. https://doi.org/10.1002/ajmg.b.32159

Greenland, S., & Robins, J. M. (1986). Identifiability, exchangeability, and epidemiological confounding. *International Journal of Epidemiology*, *15*(3), 413–419. https://doi.org/10.1093/ije/15.3.413

Grosz, M. P., Rohrer, J. M., & Thoemmes, F. (2020). The taboo against explicit causal inference in nonexperimental psychology. *Perspectives on Psychological Science*, *15*(5), 1243–1255. https://doi.org/10.1177/1745691620921521

Gruber, S., & van der Laan, M. J. (2010). An application of collaborative targeted maximum likelihood estimation in causal inference and genomics. *The International Journal of Biostatistics*, *6*(1), Article 18. https://doi.org/10.2202/1557-4679.1182

Gunasekara, F. I., Richardson, K., Carter, K., & Blakely, T. (2014). Fixed effects analysis of repeated measures data. *International Journal of Epidemiology*, *43*(1), 264–269. https://doi.org/10.1093/ije/dyt221

Gustavson, K., Ystrom, E., Stoltenberg, C., Susser, E., Surén, P., Magnus, P., Knudsen, G. P., Smith, G. D., Langley, K., Rutter, M., Aase, H., & Reichborn-Kjennerud, T. (2017). Smoking in pregnancy and child ADHD. *Pediatrics*, *139*(2), Article e20162509. https://doi.org/10.1542/peds.2016-2509

Haan, E., Sallis, H. M., Zuccolo, L., Labrecque, J., Ystrom, E., Reichborn-Kjennerud, T., Andreassen, O., Havdahl, A., & Munafò, M. R. (2021). Prenatal smoking, alcohol and caffeine exposure and ADHD risk in childhood: Parental comparisons and polygenic risk score (PRS) analyses. *MedRxiv*. Advance online publication. https://doi.org/10.1101/2021.03.25.21254087

Hannigan, L. J., Eilertsen, E. M., Gjerde, L. C., Reichborn-Kjennerud, T., Eley, T. C., Rijsdijk, F. V., Ystrom, E., & McAdams, T. A. (2018). Maternal prenatal depressive symptoms and risk for early-life psychopathology in offspring: Genetic analyses in the Norwegian Mother and Child Birth Cohort Study. *The LancetPsychiatry*, *5*(10), 808–815. https://doi.org/10.1016/S2215-0366(18)30225-6

Harrington, A. (2012). The fall of the schizophrenogenic mother. *Lancet*, *379*(9823), 1292–1293. https://doi.org/10.1016/S0140-6736(12)60546-7

Hernán, M. A. (2004). A definition of causal effect for epidemiological research. *Journal of Epidemiology and Community Health*, *58*(4), 265–271. https://doi.org/10.1136/jech.2002.006361

Hernán, M. A., & Robins, J. M. (2020). *Causal inference: What if*. Chapman & Hall/CRC.

Howe, L. J., Nivard, M. G., Morris, T. T., Hansen, A. F., Rasheed, H., Cho, Y., Chittoor, G., Ahlskog, R., Lind, P. A., Palviainen, T., van der Zee, M. D., Cheesman, R., Mangino, M., Wang, Y., Li, S., Klaric, L., Ratliff, S. M., Bielak, L. F., Nygaard, M., … the Social Science Genetic Association Consortium, & the Within Family Consortium. (2022). Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nature Genetics*, *54*(5), 581–592. https://doi.org/10.1038/s41588-022-01062-7

Hwang, L.-D., Davies, N. M., Warrington, N. M., & Evans, D. M. (2020). Integrating family-based and mendelian randomization designs. *Cold Spring Harbor Perspectives in Medicine*. Advance online publication. https://doi.org/10.1101/cshperspect.a039503

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–334. https://doi.org/10.1037/a0020761

Imbens, G. W. (2015). Matching methods in practice: Three examples. *The Journal of Human Resources*, *50*(2), 373–419. https://doi.org/10.3368/jhr.50.2.373

Jami, E. S., Eilertsen, E. M., Hammerschlag, A. R., Qiao, Z., Evans, D. M., Ystrøm, E., Bartels, M., & Middeldorp, C. M. (2020). Maternal and paternal effects on offspring internalizing problems: Results from genetic and family-based analyses. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, *183*(5), 258–267. https://doi.org/10.1002/ajmg.b.32784

Jami, E. S., Hammerschlag, A. R., Bartels, M., & Middeldorp, C. M. (2021). Parental characteristics and offspring mental health and related outcomes: A systematic review of genetically informative literature. *Translational Psychiatry*, *11*(1), Article 197. https://doi.org/10.1038/s41398-021-01300-2

Kang, H.-J., Park, Y., Yoo, K.-H., Kim, K.-T., Kim, E.-S., Kim, J.-W., Kim, S.-W., Shin, I.-S., Yoon, J.-S., Kim, J. H., & Kim, J.-M. (2020). Sex differences in the genetic architecture of depression. *Scientific Reports*, *10*(1), Article 9927. https://doi.org/10.1038/s41598-020-66672-9

Kang, J. D., & Schafer, J. L. (2007). Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical Science*, *22*(4), 523–539. https://doi.org/10.1214/07-STS227

Kendler, K. S., Ohlsson, H., Sundquist, J., & Sundquist, K. (2015). Triparental families: A new genetic-epidemiological design applied to drug abuse, alcohol use disorders, and criminal behavior in a Swedish national sample. *The American Journal of Psychiatry*, *172*(6), 553–560. https://doi.org/10.1176/appi.ajp.2014.14091127

Kendler, K. S., Ohlsson, H., Sundquist, J., & Sundquist, K. (2019). Parent-offspring transmission of drug abuse and alcohol use disorder: Application of the multiple parenting relationships design. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, *180*(4), 249–257. https://doi.org/10.1002/ajmg.b.32720

Kim, Y., Balbona, J. V., & Keller, M. C. (2021). Bias and precision of parameter estimates from models using polygenic scores to estimate environmental and genetic parental influences. *Behavior Genetics*, *51*(3), 279–288. https://doi.org/10.1007/s10519-020-10033-9

Kim, Y., & Steiner, P. (2016). Quasi-experimental designs for causal inference. *Educational Psychologist*, *51*(3–4), 395–405. https://doi.org/10.1080/00461520.2016.1207177

King, V., Pragg, B., & Lindstrom, R. (2020). Family relationships during adolescence and stepchilden's educational attainment in young adulthood. *Journal of Marriage and Family*, *82*(2), 622–638. https://doi.org/10.1111/jomf.12642

Knopik, V. S., Marceau, K., Bidwell, L. C., Palmer, R. H. C., Smith, T. F., Todorov, A., Evans, A. S., & Heath, A. C. (2016). Smoking during pregnancy and ADHD risk: A genetically informed, multiple-rater approach. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, *171*(7), 971–981. https://doi.org/10.1002/ajmg.b.32421

Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., Benonisdottir, S., Oddsson, A., Halldorsson, B. V., Masson, G., Gudbjartsson, D. F., Helgason, A., Bjornsdottir, G., Thorsteinsdottir, U., & Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science*, *359*(6374), 424–428. https://doi.org/10.1126/science.aan6877

Kuja-Halkola, R., D'Onofrio, B. M., Larsson, H., & Lichtenstein, P. (2014). Maternal smoking during pregnancy and adverse outcomes in offspring: Genetic and environmental sources of covariance. *Behavior Genetics*, *44*(5), 456–467. https://doi.org/10.1007/s10519-014-9668-4

Langley, K., Heron, J., Smith, G. D., & Thapar, A. (2012). Maternal and paternal smoking during pregnancy and risk of ADHD symptoms in offspring: Testing for intrauterine effects. *American Journal of Epidemiology*, *176*(3), 261–268. https://doi.org/10.1093/aje/kwr510

Lash, T. L., Fox, M. P., & Fink, A. K. (2009). *Applying quantitative bias analysis to epidemiologic data*. Springer. https://doi.org/10.1007/978-0-387-87959-8

Latvala, A., Kuja-Halkola, R., Långström, N., & Lichtenstein, P. (2015). Paternal antisocial behavior and sons' cognitive ability: A population-based quasiexperimental study. *Psychological Science*, *26*(1), 78–88. https://doi.org/10.1177/0956797614555726

Lawlor, D. A., Tilling, K., & Davey Smith, G. (2016). Triangulation in aetiological epidemiology. *International Journal of Epidemiology*, *45*(6), 1866–1886. https://doi.org/10.1093/ije/dyw314

Lifford, K. J., Harold, G. T., & Thapar, A. (2008). Parent-child relationships and ADHD symptoms: A longitudinal analysis. *Journal of Abnormal Child Psychology*, *36*(2), 285–296. https://doi.org/10.1007/s10802-007-9177-5

Little, R. J., & Rubin, D. B. (2000). Causal effects in clinical and epidemiological studies via potential outcomes: Concepts and analytical approaches. *Annual Review of Public Health*, *21*(1), 121–145. https://doi.org/10.1146/annurev.publhealth.21.1.121

Liu, C.-Y., Schoeler, T., Davies, N. M., Peyre, H., Lim, K.-X., Barker, E. D., Llewellyn, C., Dudbridge, F., & Pingault, J.-B. (2021). Are there causal

relationships between attention-deficit/hyperactivity disorder and body mass index? Evidence from multiple genetically informed designs. *International Journal of Epidemiology*, *50*(2), 496–509. https://doi.org/10.1093/ije/dyaa214

Loehlin, J. C. (2016). What can an adoption study tell us about the effect of prenatal environment on a trait? *Behavior Genetics*, *46*(3), 329–333. https://doi.org/10.1007/s10519-015-9730-x

Mamluk, L., Jones, T., Ijaz, S., Edwards, H. B., Savović, J., Leach, V., Moore, T. H. M., von Hinke, S., Lewis, S. J., Donovan, J. L., Lawlor, D. A., Davey Smith, G., Fraser, A., & Zuccolo, L. (2020). Evidence of detrimental effects of prenatal alcohol exposure on offspring birthweight and neurodevelopment from a systematic review of quasi-experimental studies. *International Journal of Epidemiology*, *49*(6), 1972–1995. https://doi.org/10.1093/ije/dyz272

Martens, E. P., Pestman, W. R., de Boer, A., Belitser, S. V., & Klungel, O. H. (2006). Instrumental variables: Application and limitations. *Epidemiology*, *17*(3), 260–267. https://doi.org/10.1097/01.ede.0000215160.88317.cb

Matthay, E. C., Hagan, E., Gottlieb, L. M., Tan, M. L., Vlahov, D., Adler, N. E., & Glymour, M. M. (2020). Alternative causal inference methods in population health research: Evaluating tradeoffs and triangulating evidence. *SSM—Population Health*, *10*, Article 100526. https://doi.org/10.1016/j.ssmph.2019.100526

Maurin, E., & McNally, S. (2008). Vive la révolution! Long-term educational returns of 1968 to the angry students. *Journal of Labor Economics*, *26*(1), 1–33. https://doi.org/10.1086/522071

McAdams, T. A., Hannigan, L. J., Eilertsen, E. M., Gjerde, L. C., Ystrom, E., & Rijsdijk, F. V. (2018). Revisiting the children-of-twins design: Improving existing models for the exploration of intergenerational associations. *Behavior Genetics*, *48*(5), 397–412. https://doi.org/10.1007/s10519-018-9912-4

McAdams, T. A., Rijsdijk, F. V., Neiderhiser, J. M., Narusyte, J., Shaw, D. S., Natsuaki, M. N., Spotts, E. L., Ganiban, J. M., Reiss, D., Leve, L. D., Lichtenstein, P., & Eley, T. C. (2015). The relationship between parental depressive symptoms and offspring psychopathology: Evidence from a children-of-twins study and an adoption study. *Psychological Medicine*, *45*(12), 2583–2594. https://doi.org/10.1017/S0033291715000501

McClure, E. B., Brennan, P. A., Hammen, C., & Le Brocque, R. M. (2001). Parental anxiety disorders, child anxiety disorders, and the perceived parent–child relationship in an Australian high-risk sample. *Journal of Abnormal Child Psychology*, *29*(1), 1–10. https://doi.org/10.1023/A:1005260311313

Meier, S. M., Plessen, K. J., Verhulst, F., Mors, O., Mortensen, P. B., Pedersen, C. B., & Agerbo, E. (2017). Familial confounding of the association between maternal smoking during pregnancy and internalizing disorders in offspring. *Psychological Medicine*, *47*(8), 1417–1426. https://doi.org/10.1017/S0033291716003627

Munafò, M. R., Higgins, J. P. T., & Smith, G. D. (2021). Triangulating evidence through the inclusion of genetically informed designs. *Cold Spring Harbor Perspectives in Medicine*, *11*(8), Article a040659. https://doi.org/10.1101/cshperspect.a040659

Munafò, M. R., & Smith, G. D. (2018). Robust research needs many lines of evidence. *Nature*, *553*(7689), 399–401. https://doi.org/10.1038/d41586-018-01023-3

Naimi, A. I., Cole, S. R., & Kennedy, E. H. (2017). An introduction to g methods. *International Journal of Epidemiology*, *46*(2), 756–762. https://doi.org/10.1093/ije/dyw323

Natsuaki, M. N., Shaw, D. S., Neiderhiser, J. M., Ganiban, J. M., Harold, G. T., Reiss, D., & Leve, L. D. (2014). Raised by depressed parents: Is it an environmental risk? *Clinical Child and Family Psychology Review*, *17*(4), 357–367. https://doi.org/10.1007/s10567-014-0169-z

O'Connor, T. G. (2003). Natural experiments to study the effects of early experience: Progress and limitations. *Development and Psychopathology*, *15*(4), 837–852. https://doi.org/10.1017/S0954579403000403

Obel, C., Zhu, J. L., Olsen, J., Breining, S., Li, J., Grønborg, T. K., Gissler, M., & Rutter, M. (2016). The risk of attention deficit hyperactivity disorder in children exposed to maternal smoking during pregnancy—A re-examination using a sibling design. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *57*(4), 532–537. https://doi.org/10.1111/jcpp.12478

Oldenburg, C. E., Moscoe, E., & Bärnighausen, T. (2016). Regression discontinuity for causal effect estimation in epidemiology. *Current Epidemiology Reports*, *3*(3), 233–241. https://doi.org/10.1007/s40471-016-0080-x

Parker, G. (1982). Re-searching the schizophrenogenic mother. *Journal of Nervous and Mental Disease*, *170*(8), 452–462. https://doi.org/10.1097/00005053-198208000-00003

Petersen, A. H., & Lange, T. (2020). What is the causal interpretation of sibling comparison designs? *Epidemiology*, *31*(1), 75–81. https://doi.org/10.1097/EDE.0000000000001108

Pingault, J.-B., Allegrini, A. G., Odigie, T., Frach, L., Baldwin, J. R., Rijsdijk, F., & Dudbridge, F. (2022), Research review: How to interpret associations between polygenic scores, environmental risks, and phenotypes. *Journal of Child Psychology and Psychiatry*, *63*(10), 1125–1139. https://doi.org/10.1111/jcpp.13607

Pingault, J.-B., O'Reilly, P. F., Schoeler, T., Ploubidis, G. B., Rijsdijk, F., & Dudbridge, F. (2018). Using genetic data to strengthen causal inference in observational research. *Nature Reviews Genetics*, *19*(9), 566–580. https://doi.org/10.1038/s41576-018-0020-3

Pingault, J.-B., Richmond, R., & Smith, G. D. (2022). Causal inference with genetic data: Past, present, and future. *Cold Spring Harbor Perspectives in Medicine*, *12*(3), Article a041271. https://doi.org/10.1101/cshperspect.a041271

Pingault, J.-B., Rijsdijk, F., Schoeler, T., Choi, S. W., Selzam, S., Krapohl, E., O'Reilly, P. F., & Dudbridge, F. (2021). Genetic sensitivity analysis: Adjusting for genetic confounding in epidemiological associations. *PLOS Genetics*, *17*(6), Article e1009590. https://doi.org/10.1371/journal.pgen.1009590

Pingault, J.-B., Barkhuizen, W., Wang, B., Hannigan, L. J., Eilertsen, E. M., Corfield, E., Andreassen, O. A., Ask, H., Tesli, M., Askeland, R. B., Davies Smith, G., Stoltenberg, C., Davies, N. M., Reichborn-Kjennerud, T., Ystrom, E., & Havdahl, A. (2022). Genetic nurture versus genetic transmission of risk for ADHD traits in the Norwegian Mother, Father and Child Cohort Study. *Molecular Psychiatry*. Advance online publication. https://doi.org/10.1038/s41380-022-01863-6

Platt, J. M., McLaughlin, K. A., Luedtke, A. R., Ahern, J., Kaufman, A. S., & Keyes, K. M. (2018). Targeted estimation of the relationship between childhood adversity and fluid intelligence in a US population sample of adolescents. *American Journal of Epidemiology*, *187*(7), 1456–1466. https://doi.org/10.1093/aje/kwy006

Rice, F., Harold, G. T., Boivin, J., Hay, D. F., van den Bree, M., & Thapar, A. (2009). Disentangling prenatal and inherited influences in humans with an experimental design. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(7), 2464–2467. https://doi.org/10.1073/pnas.0808798106

Robins, J. M., & Hernán, M. A. (2009). Estimation of the causal effects of time-varying exposures. In G. Fitzmaurice, M. Davidian, G. Verbeke, & G. Molenberghs (Eds.), *Longitudinal data analysis* (pp. 553–599). CRC Press.

Rodgers, J. L., Bard, D. E., Johnson, A., D'Onofrio, B., & Miller, W. B. (2008). The cross-generational mother-daughter-aunt-niece design: Establishing validity of the MDAN design with NLSY fertility variables. *Behavior Genetics*, *38*(6), 567–578. https://doi.org/10.1007/s10519-008-9225-0

Rutter, M. (2007). Proceeding from observed correlation to causal inference: The use of natural experiments. *Perspectives on Psychological Science*, *2*(4), 377–395. https://doi.org/10.1111/j.1745-6916.2007.00050.x

Sariaslan, A., Mikkonen, J., Aaltonen, M., Hiilamo, H., Martikainen, P., & Fazel, S. (2021). No causal associations between childhood family income and subsequent psychiatric disorders, substance misuse and violent crime arrests: A nationwide Finnish study of >650 000 individuals and their siblings. *International Journal of Epidemiology*, *50*(5), 1628–1638. https://doi.org/10.1093/ije/dyab099

Seeman, M. V. (2016). Schizophrenogenic mother. In J. Lebow, A. Chambers, & D. C. Breunlin (Eds.), *Encyclopedia of couple and family therapy* (pp. 1–2). Springer International Publishing. https://doi.org/10.1007/978-3-319-15877-8_482-1

Shen, C., Li, X., Li, L., & Were, M. C. (2011). Sensitivity analysis for causal inference using inverse probability weighting. *Biometrical Journal* [Biometrische Zeitschrift], *53*(5), 822–837. https://doi.org/10.1002/bimj.201100042

Shenassa, E. D., Rogers, M. L., & Buka, S. L. (2023). Maternal smoking during pregnancy, offspring smoking, adverse childhood events, and risk of major depression: A sibling design study. *Psychological Medicine*, *53*(1), 206–216. https://doi.org/10.1017/S0033291721001392

Silberg, J. L., Maes, H., & Eaves, L. J. (2010). Genetic and environmental influences on the transmission of parental depression to children's depression and conduct disturbance: An extended Children of Twins study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *51*(6), 734–744. https://doi.org/10.1111/j.1469-7610.2010.02205.x

Singh, A. L., D'Onofrio, B. M., Slutske, W. S., Turkheimer, E., Emery, R. E., Harden, K. P., Heath, A. C., Madden, P. A., Statham, D. J., & Martin, N. G. (2011). Parental depression and offspring psychopathology: A children of twins study. *Psychological Medicine*, *41*(7), 1385–1395. https://doi.org/10.1017/S0033291710002059

Singh, T., Poterba, T., Curtis, D., Akil, H., Al Eissa, M., Barchas, J. D., Bass, N., Bigdeli, T. B., Breen, G., Bromet, E. J., Buckley, P. F., Bunney, W. E., Bybjerg-Grauholm, J., Byerley, W. F., Chapman, S. B., Chen, W. J., Churchhouse, C., Craddock, N., Cusick, C. M., … Daly, M. J. (2022). Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature*, *604*(7906), 509–516. https://doi.org/10.1038/s41586-022-04556-w

Sjölander, A., Frisell, T., Kuja-Halkola, R., Öberg, S., & Zetterqvist, J. (2016). Carryover effects in sibling comparison designs. *Epidemiology*, *27*(6), 852–858. https://doi.org/10.1097/EDE.0000000000000541

Skoglund, C., Chen, Q., D'Onofrio, B. M., Lichtenstein, P., & Larsson, H. (2014). Familial confounding of the association between maternal smoking during pregnancy and ADHD in offspring. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *55*(1), 61–68. https://doi.org/10.1111/jcpp.12124

Smith, G. D., & Ebrahim, S. (2003). 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? *International Journal of Epidemiology*, *32*(1), 1–22. https://doi.org/10.1093/ije/dyg070

Thapar, A., & Rice, F. (2020). Family-based designs that disentangle inherited factors from pre- and postnatal environmental exposures: In vitro fertilization, discordant sibling pairs, maternal versus paternal comparisons, and adoption designs. *Cold Spring Harbor Perspectives in Medicine*. Advance online publication. https://doi.org/10.1101/cshperspect.a038877

Thapar, A., Rice, F., Hay, D., Boivin, J., Langley, K., van den Bree, M., Rutter, M., & Harold, G. (2009). Prenatal smoking might not cause attention-deficit/hyperactivity disorder: Evidence from a novel design. *Biological Psychiatry*, *66*(8), 722–727. https://doi.org/10.1016/j.biopsych.2009.05.032

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, *51*(6), 309–317. https://doi.org/10.1037/h0044319

Torche, F. (2018). Prenatal exposure to an acute stressor and children's cognitive outcomes. *Demography*, *55*(5), 1611–1639. https://doi.org/10.1007/s13524-018-0700-9

Tozzi, V., Lertxundi, A., Ibarluzea, J. M., & Baccini, M. (2019). Causal effects of prenatal exposure to $PM_{2.5}$ on child development and the role of unobserved confounding. *International Journal of Environmental Research and Public Health*, *16*(22), Article 4381. https://doi.org/10.3390/ijerph16224381

Trubetskoy, V., Pardiñas, A. F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T. B., Bryois, J., Chen, C.-Y., Dennison, C. A., Hall, L. S., Lam, M., Watanabe, K., Frei, O., Ge, T., Harwood, J. C., Koopmans, F., Magnusson, S., Richards, A. L., Sidorenko, J., … the Schizophrenia Working Group of the Psychiatric Genomics Consortium. (2022). Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature*, *604*(7906), 502–508. https://doi.org/10.1038/s41586-022-04434-5

van der Laan, M. J., & Rubin, D. (2006). Targeted maximum likelihood learning. *The International Journal of Biostatistics*, *2*(1), Article 11. https://doi.org/10.2202/1557-4679.1043

Van Der Wel, K. A., Östergren, O., Lundberg, O., Korhonen, K., Martikainen, P., Andersen, A. N., & Urhoj, S. K. (2019). A gold mine, but still no Klondike: Nordic register data in health inequalities research. *Scandinavian Journal of Public Health*, *47*(6), 618–630. https://doi.org/10.1177/1403494819858046

Vansteelandt, S., & Didelez, V. (2018). Improving the robustness and efficiency of covariate-adjusted linear instrumental variable estimators. *Scandinavian Journal of Statistics*, *45*(4), 941–961. https://doi.org/10.1111/sjos.12329

Wang, B., Baldwin, J. R., Schoeler, T., Cheesman, R., Barkhuizen, W., Dudbridge, F., Bann, D., Morris, T. T., & Pingault, J.-B. (2021). Robust genetic nurture effects on education: A systematic review and meta-analysis based on 38,654 families across 8 cohorts. *American Journal of Human Genetics*, *108*(9), 1780–1791. https://doi.org/10.1016/j.ajhg.2021.07.010

Yengo, L., Robinson, M. R., Keller, M. C., Kemper, K. E., Yang, Y., Trzaskowski, M., Gratten, J., Turley, P., Cesarini, D., Benjamin, D. J., Wray, N. R., Goddard, M. E., Yang, J., & Visscher, P. M. (2018). Imprint of assortative mating on the human genome. *Nature Human Behaviour*, *2*(12), 948–954. https://doi.org/10.1038/s41562-018-0476-3

Young, A. I., Benonisdottir, S., Przeworski, M., & Kong, A. (2019). Deconstructing the sources of genotype-phenotype associations in humans. *Science*, *365*(6460), 1396–1400. https://doi.org/10.1126/science.aax3710

Zetterqvist, J., Vansteelandt, S., Pawitan, Y., & Sjölander, A. (2016). Doubly robust methods for handling confounding by cluster. *Biostatistics*, *17*(2), 264–276. https://doi.org/10.1093/biostatistics/kxv041

Zhang, G., Bacelis, J., Lengyel, C., Teramo, K., Hallman, M., Helgeland, Ø., Johansson, S., Myhre, R., Sengpiel, V., Njølstad, P. R., Jacobsson, B., & Muglia, L. (2015). Assessing the causal relationship of maternal height on birth size and gestational age at birth: A mendelian randomization analysis. *PLOS Medicine*, *12*(8), Article e1001865. https://doi.org/10.1371/journal.pmed.1001865

Received March 16, 2022
Revision received October 25, 2022
Accepted December 31, 2022 ■