[21] Froehlich TE, Lanphear BP, Auinger P *et al*. Association of tobacco and lead exposures with attention-deficit/hyperactivity disorder. *Pediatrics* 2009;**124:** e1054–63.

[22] Biederman J, Monuteaux MC, Faraone SV, Mick E. Parsing the associations between prenatal exposure to nicotine and offspring psychopathology in a nonreferred sample. *J Adolesc Health* 2009;**45:**142–48.

[23] Thapar A, Rice F, Hay D, Boivin J *et al*. Response to: testing the association between smoking in pregnancy and attention-deficit/hyperactivity disorder in a novel design. *Biol Psychiatry* 2010;**68:**311–400.

[24] Gilman SE, Gardener H, Buka SL. Maternal smoking during pregnancy and children's cognitive and physical development: a causal risk factor? *Am J Epidemiol* 2008; **168:**522–31.

Published by Oxford University Press on behalf of the International Epidemiological Association

International Journal of Epidemiology 2011;**40**:345–349
doi:10.1093/ije/dyr057

Commentary: Advent of sibling designs

Stephen J Donovan[1,2] and Ezra Susser[1,3]*

[1]New York State Psychiatric Institute, New York, NY, USA [2]College of Physicians and Surgeons, New York, NY, USA and [3]Mailman School of Public Health, Columbia University, New York, NY, USA

*Corresponding author. Mailman School of Public Health, Columbia University, New York, NY 10032-3727, USA.
E-mail: ess8@columbia.edu

Accepted 25 February 2011

The rapidly increasing use of sibling comparisons is a welcome development in epidemiology. Although sibling designs have been used by epidemiologists since the mid-20th century, contemporary researchers are extending the range of applications, clarifying appropriate methods[1] and introducing novel strategies. The proliferation of sibling studies has not been matched, however, by comparable progress towards a conceptual framework that gives coherence to the new array of approaches.

We propose that, as a starting point for a conceptual framework, a fundamental distinction be made between designs that assume a 'stable' vs 'dynamic' family context. Designs that assume a stable family context exploit the fact that siblings share stable aspects of family context as well as half their genome. In 'sibling discordance' studies, for example, we compare outcomes among siblings who are discordant for an exposure of interest, and we want the siblings to be as similar as possible in family context and genetic predisposition. Differences between siblings in family context are potential confounders, and, when measured (e.g. birth order), are controlled in the analyses. Influences of family members upon one another are generally disregarded. In contrast, designs that assume a dynamic family context exploit the fact that siblings and other family members influence each other in a variety of ways. In birth-order studies, for example, we compare outcomes among siblings who encounter different family contexts and may play a role in creating these contexts. Influences of family members upon one another, and the resulting changes in family context, are the topic of investigation.

We hope to show that this distinction between sibling designs based on stable vs dynamic family context is readily understood, broadly applicable to traditional as well as novel approaches, and useful as a starting point. To this end, we portray a range of recent sibling studies, and consider them within this framework. Our overarching goal is to stimulate others to further elaborate the conceptual underpinnings of sibling designs.

Sibling designs based on stable family context

Since the sibling discordance study is probably the most widely used sibling design, we use it to characterize designs based on stable family context. The power of this approach resides in its ability to vary one aspect of the environment while keeping so much else similar. Although sibling discordance studies can be traced to the 19th century,[2] rigorous use of the design began after World War II. In the 1960s, Record *et al*.[3,4] applied it to control familial confounding in their studies of the relation of birthweight to offspring verbal reasoning scores on school examinations. This remarkable work set an early precedent

for subsequent studies[5] (as did other work by these authors).[6]

Sibling discordance studies are particularly suitable for examining the relation of prenatal exposures to offspring outcomes. We use discordance in maternal exposures or conditions across pregnancies to indicate variation in the *in utero* experience of same-sex siblings. Same-sex siblings share enough stable aspects of family environment and genetic predisposition that the strategy substantially reduces the potential for confounding. This provides a major advantage over population studies of unrelated individuals. The approach also offers a major advantage over the twin design. Twins are, by definition, products of the same pregnancy, and, while their *in utero* experience can be very different, we resort to proxies such as birth outcomes to gauge this difference.

The elegant paper by Obel *et al*.[7] in this issue of the *IJE* offers an example of a study examining a prenatal exposure that compared results from a full population to results from discordant siblings within that population. It was a registry study of a national cohort in Finland. The authors first found a strong association between maternal smoking during pregnancy and offspring ICD-10 (International Classification of Diseases, 10th Revision) Hyperkinetic Disorder in the full cohort. Next, they identified mothers who smoked during one pregnancy, but not during another, and compared hyperkinetic disorder outcomes in the relevant siblings. This second analysis found only a minimal association. They concluded that the first result was confounded by family factors (genes and/or family context) and that there was little or no causal relationship between prenatal smoking and hyperkinetic disorder. The authors noted, however, that the validity of their conclusion rested upon a crucial assumption that family factors relevant to hyperkinetic disorder were stable and therefore similar across siblings.

Notwithstanding this caveat, the most plausible interpretation is that the sibling analysis revealed that the population result was confounded. Numerous other sibling discordance studies[8–10] (as well as other approaches[11]) have also suggested that associations between prenatal smoking and offspring outcomes are especially vulnerable to uncontrolled confounding by family factors. Thus, our ability to interpret the many reported associations between prenatal smoking and offspring outcomes has gained much from the uptake of sibling discordance studies.

These studies have shown the value of conducting a sibling discordance study as a more stringent test after a potentially causal association is identified in a study of unrelated individuals. Yet, this use of sibling discordance studies has an inherent limitation. If the result is negative, suggesting that previous studies were confounded, we do not know whether the confounding was due to genes or family environment or both. Consequently, we do not gain any clear direction in the ongoing search for causes.

An intriguing approach has recently been developed that has the potential to differentiate confounding by genetic predisposition and by family context after a negative sibling discordance study. This design achieves systematic variation in genetic relatedness, which cannot be done in a sibling discordance study. The underlying logic is to examine the same prenatal exposure in different genetic contexts. The approach was made possible by the methods used in Assisted Reproductive Technology. Some of these methods create variation in the genetic relatedness of the pregnant woman to the embryo she carries.

We describe here the essence of the design and refer readers elsewhere for further detail.[12,13] The design compares the magnitude of the association between a prenatal exposure and an offspring outcome for two groups of children: genetically related to their uterine mother vs not genetically related. If the magnitude of the association is the same in the two groups of children, the result suggests confounding by family context. If the association is only present in children genetically related to their uterine mother, the result suggests confounding by genetic predisposition.

One of the first applications of this design was to investigate the relation of prenatal smoking to attention-deficit hyperactivity disorder (ADHD), which is similar to hyperkinetic disorder.[13] Among children who were genetically related to their uterine mother, prenatal exposure to maternal smoking was associated with ADHD; among children who were not, there was no association. This pattern of results suggested that genetic predisposition, rather than family context, was the main confounder of associations between prenatal smoking and ADHD. The study was small and cannot be considered conclusive, but it serves to illustrate the potential of the strategy.

Though this design is innovative and novel, it is most likely to be informative when deployed after a negative sibling discordance study has shown that a previously reported association from a population study was due to confounding. Otherwise, this design cannot distinguish a true causal effect of the exposure from confounding by family context. If a similar association between exposure and outcome were found in the two groups of children, either of these interpretations would be plausible. Also, it will produce a clear pattern of results only when either genetic predisposition or family context, but not both, are major confounders. Nonetheless, it represents an important advance, because at least in some circumstances it has the capacity to disentangle sources of confounding.

We now briefly illustrate the broadening range of applications of the sibling discordance design. It is increasingly applied to address questions that

have emerged from the use of new technologies.[13–15] In population studies, assisted fertilization is associated with adverse birth outcomes (e.g. lower birthweight and shorter gestation), and it has been postulated, quite plausibly, that these associations are due to the techniques used in assisted fertilization. Similar to the study by Obel *et al.*,[7] a recent publication compared results from a full population to results from discordant siblings within that population.[14] In this instance, the siblings were discordant in being conceived with or without assisted fertilization. The authors found that assisted fertilization was associated with adverse birth outcomes in the full population, but not in the analysis of discordant sibling pairs. They inferred that the association of assisted fertilization with adverse outcomes could be due to stable characteristics of the couple, such as factors underlying infertility, rather than due to the assisted fertilization.

A second example is a study in which siblings discordant for periconceptual exposure to famine were used to examine whether this exposure had epigenetic effects that persisted until follow-up in midlife.[15] The study was small, but the results were consistent with an epigenetic effect on insulin-like growth factor 2 (IGF2) that was hypothesized a priori. Here, the discordant siblings were used to create a more compelling design within a small sample, rather than to test the validity of a previous result from a population study. Given the scant knowledge of the full range of epigenetic variation within populations, and of the causes of such variation, a study of discordant siblings was more likely to produce a meaningful result, by partially controlling a host of unknown and unmeasured family and genetic factors that were potential confounders.

As a segué to the next section, we consider a longstanding sibling design of a different ilk. The younger siblings of a child with a relatively rare disorder (e.g. $\leqslant 1\%$ of children) are studied prospectively to identify causes of the disorder. If younger siblings have a much higher risk of the disorder than unrelated individuals (which they often do), the approach makes it feasible to conduct an in-depth prospective study on a reasonable scale. Comparisons are usually made across rather than within families; the sibling with the disorder serves as a marker for high risk. A current example is a study that is following younger siblings of children with autism.[17] Although these studies tend to assume a stable family context, they could also be used to examine effects of a dynamic family context.[18] Still another variation in this theme is a randomized controlled trial that targets the younger siblings of a child with the disorder, as was done in a trial of folic acid supplementation among mothers who previously had a child with a neural tube defect.[19] The implicit assumption is a stable family context. The randomization, however, may obviate the need to measure family context.

Sibling designs based on dynamic family context

From a broad array of sibling designs based on dynamic family context, we select two to characterize it, and a third to indicate the potential for new designs of this kind. We begin with birth order, which has known effects on health (e.g. birthweight). The birth of a sibling immediately changes the family social environment, and sets in motion a dynamic process that is manifest in varying relationships between siblings over the course of development. The birth of a sibling also follows on the changes introduced by the birth of previous siblings, which include the biological state of the mother.

The investigation of birth order and intelligence quotient (IQ) provides a good example of both the continuing value of earlier studies and of how recent studies have made significant advances on them. The notion that birth order is related to intelligence was put forward by Galton in the 19th century, but the first rigorous studies were done in the 1970s. A particularly influential series of studies was based in large cohorts of 18-year-old males in The Netherlands who were administered Raven Progressive Matrices and other neuropsychological tests at The time of military induction.[20–22] These were not initially sibling studies, but comparisons of individuals with different birth orders, controlling for other factors such as family size and interpregnancy interval. The last study in the series, however, did include a comparison between male siblings.[22] These studies showed (among other things) that: (i) higher birth order was related to higher scores on the Ravens and other tests; (ii) the relation of birth order to intelligence differed from the relation of birth order to height, which was taken as an indication that social aspects of family environment were involved in the relation to intelligence; and (iii) the relation held for comparisons between brothers in two-child families. To keep these findings in perspective, we note that the influence of birth order on IQ was much smaller than that of social class.

These findings were subsequently disputed on many grounds. A recent series of studies among Norwegian conscripts, however, replicated and extended them in a more definitive design.[23–25] Similar to the earlier Dutch studies, the Norwegian studies suggested that higher birth order is associated with higher IQ, that social rather than biological rank in the family accounts for the effect, and that the findings hold for within, as well as between, family comparisons. One of the ways in which the recent studies differed was that they explicitly included within, as well as between, family comparisons from the outset. These investigators also introduced an ingenious extra element in the design. They showed that boys who became the eldest after the death of their elder brother resembled the first born in the population,

not the second born; and, similarly, that, boys who became the second eldest resembled the second born in the population, not the third born. Based on this comparison, they could infer that the birth order effect on intelligence reflected social position within the family.

Thus, it is now clear that the interpretation of the relationship between birth order and intelligence requires consideration of the family configuration as a dynamic social process.[26] Although this topic is beyond our scope here, we note two intriguing findings that underscore the importance of this perspective. First, children without siblings have, on average, a lower IQ than the eldest sibling. Secondly, the advantage of the elder sibling does not appear to be constant over time, but rather, varies over the course of development, and, according to some studies, may not be manifest until adolescence.

For a second example, we consider the interpregnancy interval. Like birth order, an effect of the spacing of pregnancies could reflect an influence on the family configuration, as well as on the biological state of the mother. An effect of very short interpregnancy interval, in particular, may be plausibly related to the latter if a pregnancy biologically depletes a mother in some way: the shorter the recovery, the greater the residual depletion. The early series of studies on birth order and IQ described above were, to our knowledge, the first to rigorously examine the impact of interpregnancy interval within families (they found no effect on IQ).[22]

Recently, a large study carried out in California examined risk of autism in a second sibling as a function of interpregnancy interval.[27] The risk for autism in the second sibling was increased after a short interpregnancy interval, and especially after a very short interpregnancy interval (<12 months). The authors noted that a possible interpretation was an alteration of maternal physiology, perhaps due to maternal nutritional depletion. In this scenario, the birth of the first sibling influences maternal biology, and thereby the health of the second sibling. The influence of one birth upon another is the topic of interest.

A third design based on dynamic family context is still being developed but points to future directions. We have dubbed it 'the older/younger sibling design'. It involves an older sibling with an outcome (e.g. substance use disorder) who is considered to be not only a marker of increased risk but also a direct influence on a younger sibling. We are developing this design so as to consider the age-specific family environment that the younger sibling lives in as well as the larger social (peer/neighbourhood) environment the younger sibling must enter. Thus, it could offer a means to study the social processes by which high-risk children and adolescents progress to develop a health outcome, with implications for intervention.

Conclusion

Every complete medical exam elicits a family history, which influences clinical decisions. The advent of sibling designs in epidemiology reflects an awareness of the tremendous amount of information that can be obtained from examining health outcomes within the context of the family. Sibling designs, both stable and dynamic, have great potential for further development, and offer an avenue for significant advance in epidemiology.

Funding

This work was partly supported by the National Institute on Aging, of the National Institutes of Health (P01 AG 023028); Ezra Susser, Principal Investigator.

Conflict of interest: None declared.

References

[1] Susser E, Eide MG, Begg M. Invited commentary: The use of sibship studies to detect familial confounding. *Am J Epidemiol* 2010;**172:**537–39.

[2] Sullivan WC. A note on the influence of maternal inebriety on the offspring. *J Ment Sci* 1899;**45:**489–503 [reprinted *Int J Epidemiol* 2011;**40**:278–82.

[3] Record RG, McKeown T, Edwards JH. The relation of measured intelligence to birth weight and duration of gestation. *Ann Hum Genet* 1969;**33:**71–79.

[4] Record RG, McKeown T, Edwards JH. An investigation of the difference in measured intelligence between twins and single births. *Ann Hum Genet* 1970;**34:**11–20.

[5] Paneth N, Susser E, Susser M. Origins and early development of the case-control study: Part 2, The case-control study from Lane-Claypon to 1950. *Soz Praventivmed* 2002; **47:**359–65.

[6] Matte TD, Bresnahan M, Begg MD, Susser E. Influence of variation in birth weight within normal range and within sibships on IQ at age 7 years: cohort study. *BMJ* 2001; **323:**310–14.

[7] Obel C, Olsen J, Henriksen TB *et al*. Is maternal smoking during pregnancy a risk factor for Hyperkinetic disorder? Findings from a sibling design. *Int J Epidemiol* 2011;**40:** 338–45.

[8] Knopik VS. Commentary: smoking during pregnancy–genes and environment weigh in. *Int J Epidemiol* 2010; **39:**1203–5.

[9] Iliadou AN, Koupil I, Villamor E *et al*. Familial factors confound the association between maternal smoking during pregnancy and young adult offspring overweight. *Int J Epidemiol* 2010;**39:**1193–202.

[10] Gilman SE, Gardener H, Buka SL. Maternal smoking during pregnancy and children's cognitive and physical development: a causal risk factor? *Am J Epidemiol* 2008; **168:**522–31.

[11] Breslau N, Paneth N, Lucia VC, Paneth-Pollak R. Maternal smoking during pregnancy and offspring IQ. *Int J Epidemiol* 2005;**34:**1047–53.

[12] Thapar A, Harold G, Rice F *et al*. Do intrauterine or genetic influences explain the foetal origins of chronic disease? A novel experimental method for disentangling effects. *BMC Med Res Methodol* 2007;**7:**25.

[13] Thapar A, Rice F, Hay D *et al*. Prenatal smoking might not cause attention-deficit/hyperactivity disorder: evidence from a novel design. *Biol Psychiatry* 2009;**66:**722–27.

[14] Sutcliffe AG, Ludwig M. Outcome of assisted reproduction. *Lancet* 2007;**370:**351–59.

[15] Heijmans BT, Tobi EW, Stein AD *et al*. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc Natl Acad Sci USA* 2008;**105:** 17046–49.

[16] Sun L. Early childhood general anaesthesia exposure and neurocognitive development. *Br J Anaesth* 2010; **105(Suppl 1):**i61–i68.

[17] Zwaigenbaum L, Thurm A, Stone W *et al*. Studying the emergence of autism spectrum disorders in high-risk infants: methodological and practical issues. *J Autism Dev Disord* 2007;**37:**466–80.

[18] Ozonoff S, Young GS, Steinfeld MB *et al*. How early do parent concerns predict later autism diagnosis? *J Dev Behav Pediatr* 2009;**30:**367–75.

[19] MRC Vitamin Study Research Group. Prevention of neural tube defects: results of the Medical Research Council Vitamin Study. *Lancet* 1991;**338:**131–37.

[20] Belmont L, Marolla FA. Birth order, family size, and intelligence. *Science* 1973;**182:**1096–101.

[21] Belmont L, Stein ZA, Susser MW. Comparison of associations of birth order with intelligence test score and height. *Nature* 1975;**255:**54–56.

[22] Belmont L, Stein Z, Zybert P. Child spacing and birth order: Effect on intellectual ability in two-child families. *Science* 1978;**202:**995–96.

[23] Bjerkedal T, Kristensen P, Skjeret GA, Brevik JI. Intelligence test scores and birth order among young Norwegian men (conscripts) analyzed within and between families. *Intelligence* 2007;**35:**503–14.

[24] Kristensen P, Bjerkedal T. Explaining the relation between birth order and intelligence. *Science* 2007;**316:**1717.

[25] Sulloway FJ. Birth order and intelligence. *Science* 2007; **316:**1711–12.

[26] Zajonc RB, Sulloway FJ. The confluence model: birth order as a within-family or between-family dynamic? *Pers Soc Psychol Bull* 2007;**33:**1187–94.

[27] Cheslack-Postava K, Liu K, Bearman PS. Closely spaced pregnancies are associated with increased odds of autism in California sibling births. *Pediatrics* 2011;**127:**246–53.