

5. Meyer M, Tonascia J. Maternal smoking, pregnancy complications, and perinatal mortality. *Am J Obstet Gynecol* 1977;128:494–502.
6. Wilcox AJ. Birthweight and perinatal mortality: the effect of maternal smoking. *Am J Epidemiol* 1993;137:1098–104.
7. Joseph KS, Demissie K, Platt RW, Ananth CV, McCarthy BJ, Kramer MS. A parsimonious explanation for intersecting perinatal mortality curves: understanding the effects of race and of maternal smoking. *BMC Pregnancy Childbirth* 2004;4:7.
8. Smoking and pregnancy: is the question answered? Tobacco Institute Inc., 1971. Legacy Tobacco Documents Library. <http://legacy.library.uscf.edu/tid/cyw36b00> (29 July 2014, date last accessed).
9. Hickey RJ, Clelland RC, Bowers EJ. Maternal smoking, birth weight, infant death, and the self-selection problem. *Am J Obstet Gynecol* 1978;131:805–11.
10. Jones P, presentation to the Oklahoma Association of Tobacco Distributors Convention, 28 March 1981. Legacy Tobacco Documents Library. <http://legacy.library.uscf.edu/tid/kke91f00> (29 July 2014, date last accessed).
11. Lilienfeld DE. Abe and Yak: the interactions of Abraham M. Lilienfeld and Jacob Yerushalmy in the development of modern epidemiology (1945–1973). *Epidemiology* 2007;18:507–14.
12. Untitled letter, 6 March 1993. Legacy Tobacco Documents Library. <http://legacy.library.uscf.edu/tid/csw5aa00> (29 July 2014, date last accessed).
13. Hernán MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15:615–25.
14. Joseph KS, Liu S, Demissie K *et al.*, for the Fetal and Infant Health Study Group of the Canadian Perinatal Surveillance System. A parsimonious explanation for intersecting perinatal mortality curves: understanding the effect of plurality and of parity. *BMC Pregnancy Childbirth* 2003;3:3.
15. American Academy of Pediatrics Committee on Fetus and Newborn. Nomenclature for duration of gestation, birth weight and intra-uterine growth. *Pediatrics* 1967;39:935–39.
16. Kramer M. Determinants of low birth weight: methodological assessment and meta-analysis. *Bull World Health Organ* 1987;65:663–737.
17. Kramer MS, Zhang X, Platt RW. Analyzing risks of adverse pregnancy outcomes. *Am J Epidemiol* 2013;179:361–67.
18. Kramer MS, Zhang X, Platt RW. Kramer *et al.* respond to ‘Composite outcomes and paradoxes’. *Am J Epidemiol* 2013;179:371–72.
19. Yudkin PL, Wood L, Redman CWG. Risk of unexplained stillbirth at different gestational ages. *Lancet* 1987;i:1192–94.
20. Pearl J. Causal diagrams for empirical research. *Biometrika* 1995;82:669–88.
21. Hernandez-Diaz S, Schisterman EF, Hernan MA. The birth weight ‘paradox’ uncovered? *Am J Epidemiol* 2006;164:1115–20.
22. VanderWeele TJ, Mumford S, Schisterman E. Conditioning on intermediates in perinatal epidemiology. *Epidemiology* 2012;23:1–9.
23. Basso O, Wilcox AJ. Intersecting birth weight-specific mortality curves: solving the riddle. *Am J Epidemiol* 2009;169:787–97.

## Commentary: Smoking in pregnancy and offspring health: early insights into family-based and ‘negative control’ studies?

International Journal of Epidemiology, 2014, 1381–1388

doi: 10.1093/ije/dyu166



Katherine M Keyes,<sup>1,2\*</sup> George Davey Smith<sup>3</sup> and Ezra Susser<sup>1,2</sup>

<sup>1</sup>Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, USA,

<sup>2</sup>New York State Psychiatric Institute, New York, USA and <sup>3</sup>MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

\*Corresponding author. Department of Epidemiology, Mailman School of Public Health, 722 West 168th Street, #503, New York 10032, USA. E-mail: [kmk2104@columbia.edu](mailto:kmk2104@columbia.edu)

The seminal 1971 paper by Jacob Yerushalmy,<sup>1</sup> reprinted in this issue of *IJE*, was one of the earliest attempts to systematically address the problems of inferring causality from associations between prenatal maternal smoking and offspring outcomes. It falls within the broader context of his contributions to causal inference from observational

studies in the 1950s and 1960s.<sup>2–4</sup> It goes beyond his earlier work, in that he takes on a set of findings concerning the relation between prenatal cigarette smoking and neonatal mortality that were extremely puzzling at the time, and indeed remain so. Yerushalmy documents the now well-replicated finding that low birthweight offspring of

smokers have mortality advantages compared with low birthweight offspring of non-smokers, and poses questions about how this could be explained.

Setting the stage, the reprinted paper is also noteworthy as one of the earliest reports on the Child Health and Development Study (CHDS). This was a large pregnancy/birth cohort established by Yerushalmy in Oakland, California, at about the same time that the multisite Collaborative Perinatal Project (CPP) was established across 13 sites in the USA (both cohorts were born mainly in 1959–66).<sup>5–7</sup> In addition to their large size and detailed collection of prenatal data, a distinct and innovative feature shared by these two cohorts was the collection and archiving of biological specimens, including prenatal maternal and cord serum. The cohorts are now grown men and women in their 50s, and subsamples have been studied throughout critical points in their lives. They have provided the platform for a plethora of investigations, usually as separate cohorts but occasionally combined together.<sup>5</sup> The potential effects of maternal smoking have figured prominently in studies of both cohorts. In the CHDS, for example, maternal smoking during pregnancy has been associated with a range of adverse outcomes, including adverse growth and neurocognition in childhood,<sup>8,9</sup> hyperactivity in adolescence,<sup>10</sup> bipolar disorder<sup>11</sup> and lower mammographic density in adulthood.<sup>12</sup> Determining whether these associations reflect causal relations, however, remains elusive.

We focus here on the issue that Yerushalmy believed was at the heart of the puzzle he presents. He argued that potential bias in the observational evidence arose from non-comparability between maternal smokers and non-smokers. Crucially, this included unmeasured aspects of the family context.

A very large number of studies have now documented associations of prenatal maternal smoking with adverse offspring health outcomes over the life course,<sup>13</sup> but with few exceptions (e.g. low birthweight,<sup>14</sup> sudden infant death syndrome<sup>15</sup>), it remains unclear whether these associations reflect causal relationships. Yerushalmy had previously been a harsh critic of causal inference from ‘ecologic’ studies that lacked information on individuals.<sup>16,17</sup> He went on to argue that it is perilous to ignore information at the group level, in this instance the family context. More than 40 years later, evaluating the role of family context and isolating prenatal exposures as potential causes of offspring health remains a central challenge for epidemiological investigations. At the time Yerushalmy’s paper was published, the potential for bias due to absence of family contextual data was already recognized,<sup>18</sup> but it had received less attention and was less understood than ‘ecologic’ bias due to absence of individual data, and it had

rarely been discussed in relation to prenatal smoking with the depth that Yerushalmy offers. To our knowledge the distinct contributions of information from various levels of organization (e.g. from individual to family to society) was not well elaborated in epidemiology until after Yerushalmy’s paper.<sup>19</sup>

It is worth emphasizing that Yerushalmy titled the paper ‘The relationship of parents’ cigarette smoking to outcome of pregnancy...’. Despite the fact that smoking was more common and less socially sanctioned during the early 1960s when the women of the CHDS and their husbands were recruited, there were many differences between smokers and nonsmokers, both women and their husbands. Among other differences, the pregnant smokers were less likely to use contraceptive methods and less likely to have planned the pregnancy, and both smoking women and smoking husbands were more likely to drink more alcohol as well as coffee. These analyses are additionally perplexing, however, as the highest infant death rate among low birthweight infants was observed among those families in which the husband smoked and the mother did not, whereas the death rate was lowest among those families in which the mother smoked and the husband did not. Although Yerushalmy struggled to find any good explanation for this pattern of results, he used them to make a compelling case that factors associated with the family context could exert a strong bias on pregnancy studies.

Yerushalmy concludes that ‘What is needed... is to develop auxiliary and complementary methods which would overcome their built in limitations’. Because every design has unique biases and sources of non-comparability across groups, consensus across a variety of studies with different designs is necessary rather than interpretation of any one particular study. The importance of triangulating evidence, through rigorous comparison across designs each with different potential sets of biases, has long been recognized in epidemiology as a critical way in which to improve causal inference.<sup>20</sup> In the ensuing decades, epidemiologists and other health researchers have taken up this challenge, designing and improving novel studies and harnessing strategic comparisons in order to move past non-comparability towards a more robust understanding of the offspring health consequences of maternal smoking.

In the literature regarding smoking in pregnancy, various family-based designs have been proliferating as a way to achieve a more robust inference.<sup>18,21</sup> This is not a new strategy; family-based designs to understand prenatal influences on offspring health have been documented since the 19th century (see Sullivan, 1899, reprinted in *IJE* in 2011<sup>22</sup>). The reasons to use a family-based strategy are 2-fold. First, there may be sources of genetic

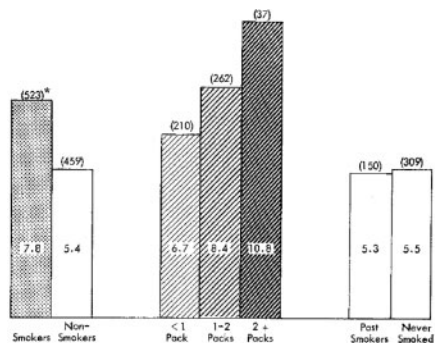


Fig. 1. Per cent of infant weighing less than 5 1/2 lbs. according to the cigarette smoking habits of their fathers. Includes smokers of unknown number of packs.

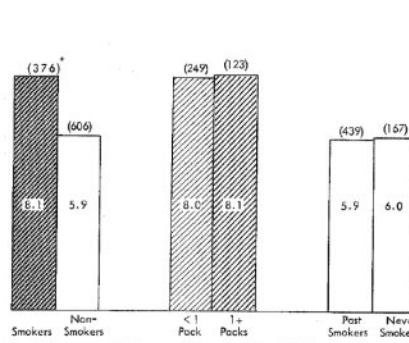


Fig. 2. Per cent of infants weighing less than 5 1/2 lbs. according to the cigarette smoking habits of their mothers. Includes smokers of unknown number of packs.

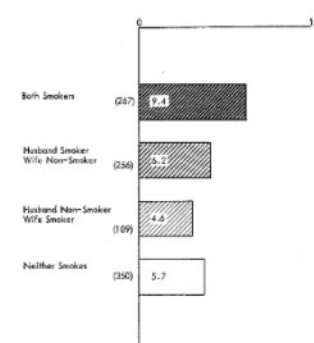


Fig. 3. Per cent of infants weighing less than 5 1/2 lbs. according to the cigarette smoking habits of the husband and wife.

Figure 1. Reprint\* of Yerushalmy's 1962 paper Figures 1 through 3. Figure 1 shows the percent low birthweight according to smoking status of the father. Figure 2 shows the percent low birthweight according to smoking status of the mother. Figure 3 shows percent low birthweight according to the smoking status of both mother and father.

\*Reprinted from Yerushalmy J. Statistical considerations and evaluation of epidemiological evidence. In: James G, Rosenthal T (eds). *Tobacco and Health*. Springfield, IL: Charles C. Thomas, 1962.

non-comparability across families. For example, there is evidence for shared genetic variance influencing both smoking and hyperactivity.<sup>23-25</sup> Thus, women who smoke in pregnancy may also carry a higher risk of hyperactivity themselves, and may pass on genetic risk factors for hyperactivity to offspring. Second, other sources of non-comparability that are shared within a family, such as the micro-family rearing environment, and subtleties of social position, are difficult to measure and fully control in standard analytical approaches. In some contexts, control for maternal cognition and education seem to be adequate to adjust for family-level factors,<sup>26-28</sup> but it is impossible to know whether such adjustments are fully sufficient to eliminate non-comparability. Although we still do not have definitive strategies for untangling the underlying causal relationships, at least within a single observational study, various approaches taken together are often complementary, as Yerushalmy suggested. When statistically-adjusted results triangulate with those estimated from other designs (e.g. Mendelian randomization<sup>29,30</sup> or between-sibling comparisons), we can be more confident that the statistical controls are sufficient for valid inference. For instance in the case of maternal smoking and offspring birthweight, discussed by Yerushalmy, the current data using a variety of methods suggest that the effect is causal. This includes randomized controlled trials,<sup>31-33</sup> non-genetic instrumental variables analyses,<sup>34</sup> Mendelian randomization studies<sup>35</sup> and maternal-paternal comparisons.<sup>21</sup>

Yerushalmy was among the first that we can find in the literature to use paternal smoking status as a 'negative control', capturing aspects of the family context to study offspring health,<sup>21,36-38</sup> with his first documented use of the method in a 1962 publication from an early group of the

CHDS participants (reprinted in Figures 1 and 2).<sup>4</sup> The logic of the maternal-paternal comparison strategy begins with the contention that the father's smoking should have less direct impact on the fetus compared with direct exposure from the mother. Similarly to maternal smoking, however, the family context and germline genetic variants are likely to confound the relation between the father's smoking and offspring health. Therefore, the relation between paternal smoking and offspring health can be used as a proxy for the amount of confounding based on shared family and genetic factors. The validity of the design rests on: (i) shared family factors are associated with maternal and paternal smoking to the same degree; (ii) germline genetic variants associated with smoking are in the same way in mothers and fathers; and (iii) no unmeasured factors confound the relation between maternal smoking and offspring health but not paternal smoking. As a hypothetical example of the latter, consider the association of parental smoking with offspring hyperactivity. Parental smoking may be similarly associated with parental cognitive ability for mothers and fathers, but maternal cognitive ability may have a unique influence on behavioral development (because, for example, the mother spends more time with the offspring and thus has a stronger influence). The negative control approach would not account for this aspect of confounding, and a stronger association of maternal smoking with offspring hyperactivity than for paternal smoking would be observed.

In Figures 1 and 2 we reprint the 1962 publication findings, as they demonstrate the clarity of the maternal-paternal comparison approach and illustrate the difficulty of 'self-selection', as Yerushalmy terms it,<sup>39</sup> in interpreting the results. Yerushalmy documents that the probability of low birthweight (defined in his analysis as

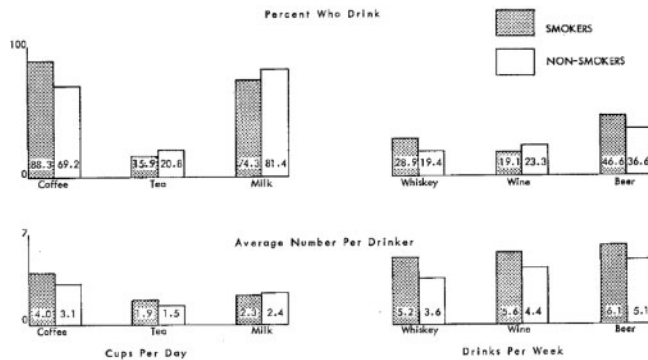


Fig. 4. Per cent who drink various beverages and the average number of drinks per drinker for smoking and non-smoking husbands.

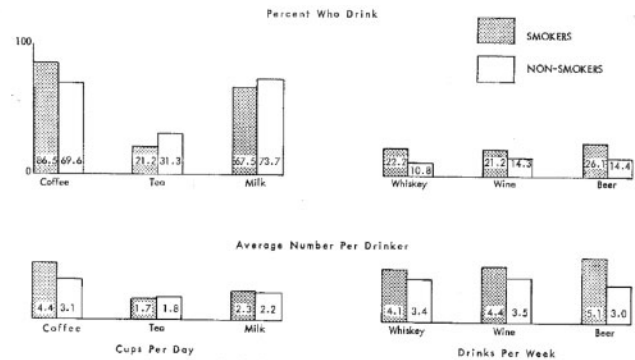


Fig. 5. Per cent who drink various beverages and the average number of drinks per drinker for smoking and non-smoking wives.

**Figure 2.** Reprint\* of Yerushalmy's 1962 paper Figures 4 and 5, showing differences between smokers and nonsmokers.

\*Reprinted from Yerushalmy J. Statistical considerations and evaluation of epidemiological evidence. In: James G, Rosenthal T (eds). *Tobacco and Health*. Springfield, IL: Charles C. Thomas, 1962.

weighing less than 5.5 pounds) offspring increases if the father smokes compared with if the father does not smoke, and that the magnitude of the relation between father smoking and offspring birthweight is similar to if not stronger than the relation between mother smoking and offspring birthweight. Yerushalmy suggests this as evidence that the relation between smoking and birthweight is unlikely to be causal. Although we now know that this conclusion is quite wrong, the logic upon which it is based remains prescient. Curiously, he also reports that birthweight was lower when both the mother and father smoked during the pregnancy period, but not when only one parent smoked, compared with no parent smoking.<sup>4</sup> It is unclear why these data led to what we now know is a wrong answer about smoking and birthweight; thus whereas he was the first to capitalize on the within-family comparison as a gauge of family-level confounding, the early papers remain a cautionary tale regarding too much belief in the data without multiple confirmatory methods and sources, which as mentioned above now strongly support the notion that maternal smoking causally influences offspring birthweight. MacMahon and colleagues used the same maternal-paternal comparison strategy in a 1966 paper<sup>40</sup> but counter Yerushalmy's findings, demonstrating little association of paternal smoking with birthweight compared with a strong effect of maternal smoking. There is now a wealth of evidence confirming the MacMahon *et al.* account.<sup>10,21,36</sup> MacMahon *et al.* and Yerushalmy's investigations also differed on other findings, underscoring the necessity of basing conclusions on the result of the weight of evidence from multiple studies with diverse designs, settings and sample characteristics. Even an association that we now take for granted as causal—pregnancy smoking and birthweight—can vary substantially from one study to another.

Another increasingly common family-based design is to compare siblings from the same biological mother across two pregnancies, one in which the woman smoked and one in which she did not.<sup>18</sup> Discordant sibling designs have generally found no effect or little effect of smoking in pregnancy on other offspring outcomes that have been associated with prenatal smoking, including hyperactivity and other behavioural problems,<sup>41–43</sup> cognition,<sup>42</sup> body size<sup>44</sup> and blood pressure.<sup>45</sup> Decreases in the effect between full cohort comparisons and discordant sibling comparisons certainly suggest that there is strong shared stable familial confounding. However, this reduction in the effect may also be attributable, to a currently unknown degree, to methodological artefacts such as increased misclassification in the discordant design, rather than amount of shared familial confounding.<sup>46,47</sup> Indeed, discordant sibling studies confirm well-documented effects of prenatal smoking on lower offspring birthweight,<sup>48</sup> but effect estimates are lower than have been documented in the literature.<sup>49</sup> This could reflect bias due to increased misclassification—e.g. in cases where mothers apparently did not smoke in one pregnancy and did in a subsequent one, the first pregnancy data may be more likely to reflect the misreporting or misrecording of the data. Finally, the contribution of unshared family environment is likely to be substantial for many health outcomes of interest;<sup>50,51</sup> in these circumstances, the use of a sibling design may not be the most efficient approach.

Thus, complementary methods to round out 'paternal negative control' and 'discordant sibling' designs must be undertaken to triangulate evidence, such as Mendelian randomization or other novel observational designs.<sup>29,52</sup> Along these lines, strategic comparisons that separate family context from prenatal exposures have recently been explored using genetically informative designs with



offspring who are either genetically or not genetically related to the woman who carried them *in utero* or to the parents who reared them. Specifically, children genetically unrelated to the carrier are often not genetically related to women who carry them *in utero*, and are also not reared by them. Therefore, the association between health behaviours in pregnancy and offspring health outcomes is not confounded by genetic vulnerability or factors associated with rearing. In some early attempts at such designs, one study reported that, among those offspring genetically unrelated to the carrier, smoking exposure *in utero* was not related to offspring hyperactivity at mean age 8 years,<sup>25</sup> whereas another analysis reported that it was related to offspring conduct disorder symptoms at mean age 5 years.<sup>53</sup> Such designs are now being further developed and applied in larger studies and to a wider range of outcomes. The separation of potential genetic and environmental confounding can also be extended using adoptive family designs. For example, children adopted into non-relative families are not genetically related to the parents who rear them; when data are available on the birth mother, the effect of prenatal exposure can be separated from postnatal rearing environment. Gaysina *et al.*<sup>53</sup> demonstrated a relation between prenatal smoking exposure and offspring conduct disorder at mean age 5 years among such a sample, strengthening the evidence for a causal relation between these factors.

Though much of the innovation in epidemiological study design and analysis is in controlling for aspects of the shared family context, it is worth noting that family circumstances are often not stable, but dynamic in nature.<sup>18,51</sup> For example, family socioeconomic status, rearing practices and maternal health behaviours are not necessarily fixed characteristics within families across time. Even among the fixed characteristics of families, the effect of these characteristics may differ across siblings (for example, two siblings may be exposed to parental divorce, but the effects of the divorce on mental health of the child could depend on personality traits and age, among other factors),<sup>50,51,54</sup> suggesting that such sibling characteristics may modify the effect of shared family factors.

A well-known modifier of family effects is birth order. Firstborn children have, on average, higher intelligence than later-born children. The causal mechanism underlying this finding was debated for some time, given that it was unclear whether there was a social mechanism (e.g. parents give more attention to the firstborn) or biological mechanism (e.g. maternal antibodies change across pregnancies). Kristensen and Bjerkedal,<sup>55</sup> however, found strong evidence of a social mechanism, in that when 'social' birth order changes (such as when a biologically second-born child becomes a social firstborn if the older sibling dies), the relation

between birth order and intelligence changes to correspond to this altered social birth order. This suggests a strong role of a dynamic family context in explaining the relation between birth order and intelligence. Such effects are potentially important for understanding individual differences in outcomes within families.

As rich data sources with observed measures of time-varying 'modes-of-life' become increasingly available, innovative and creative strategies to incorporate dynamic exposures within and across families into epidemiological studies of prenatal smoking exposure will undoubtedly inform this growing literature. Analytical methods to capture time-varying confounding are increasingly well developed for use in observational epidemiological studies,<sup>57,58</sup> including rigorous analytical methods to account for complex longitudinal causal structures in which variables can play different roles (e.g. confounders, mediators, colliders) depending on the time point and other variables in the causal model. However, our ability to statistically adjust for confounding remains limited. Strategic comparisons that illustrate the amount of confounding that is present, rather than approaches that control confounding away, remain a critical tool in assessing the strength of a potential epidemiological relation. For example, MacMahon's 1966 study suggests that among offspring of women who smoked the same amount during pregnancy, those offspring of women who smoked more prior to pregnancy than the pregnancy amount had offspring with higher birthweights compared with women who did not change smoking habits.<sup>40</sup> In 1972, Yerushalmy showed that the incidence of low birthweight infants was higher among women who became smokers after childbirth than among lifelong non-smokers, and was similar to the incidence of low birthweight offspring among women who smoked during pregnancy.<sup>56</sup> This certainly suggests that characteristics of women who change smoking habits for the better during pregnancy may be different from women who continue to smoke in ways that promote positive neonatal health.

Stable and/or shared as well as dynamic and/or unshared elements of the family context will eventually be modelled more effectively, combining methods across study designs into increasingly sophisticated approaches to understanding health across the life course. However, statistical approaches may always be relatively limited, and we note that not all variance over time may be predictable; stochastic factors may explain a substantial proportion of unshared elements of the family context or 'non-shared environment' as termed by Plomin (see Plomin, 1987,<sup>50</sup> reprinted in *IJE* with commentaries in 2011).<sup>59-63</sup> These stochastic elements can be adequately modeled.<sup>51</sup>

One interpretation of Yerushalmy's finding that, at any birthweight, infant mortality was lower if the mother was

a smoker rather than a non-smoker is that causes of lower birthweight other than smoking may have more detrimental influences on outcomes, such as infant mortality, than smoking does. Indeed, a similar picture is seen for other factors related to birthweight. For instance, maternal height is associated with higher birthweight, but at a given birthweight some infant outcomes are less favourable with greater maternal height, as Yerushalmy showed elsewhere.<sup>39</sup> Similarly, in sibling comparisons, higher birthweight of a sibling is, on average, associated with higher own birthweight but, at any given birthweight, perinatal mortality is higher if the sibling has a higher birthweight.<sup>64</sup> Whereas unmeasured confounding could potentially explain all of these relations, the fact is that such an unmeasured factor would need to be unassociated with maternal smoking, maternal height or sibling birthweight. It is difficult to conceptualize such a factor. One potential explanation, however, is that this may reflect stochastic (and by definition unsystematic) processes that lower birthweight and increase infant mortality.<sup>51</sup> MacMahon and colleagues noted this potential explanation in their 1966 analysis of birthweight and mortality,<sup>40</sup> considering the case of offspring sex. Female offspring have lower average birthweight than male offspring, but the factor that shifted the weight downward (sex) does not increase mortality. Thus, to quote MacMahon *et al.*:

...for the purpose of predicting mortality from birthweight, specification in terms of absolute weight does not mean the same for male infants as for females; at any given weight the infant in the series with lower mean weight (females) will have, relative to males, a smaller proportion of members whose weight is reduced by those factors that are associated with increased mortality, and the group will consequently have a more favorable mortality rate. Similarly, for the offspring of smokers, if their weight is reduced but their over-all mortality unaffected, then at any given weight they will, relative to the offspring of nonsmokers, have a lower component of infants premature by gestation, and consequently, a more favorable mortality rate.

We interpret MacMahon and colleagues as suggesting that if some exposures decrease birthweight but are not associated with increased mortality, the group with lower birthweight will always show a mortality advantage at any particular weight because the factors (potentially stochastic) which both lower birthweight and increase mortality will be less prevalent among the group with a known reason for their lower average birthweight.

Yerushalmy was wrong in many of his conclusions about the health effects of smoking, but he was right that 'complementary and auxiliary methods are needed to

understand these effects, because each study design has unique biases that, alone, cannot be sufficient for strong inference'.<sup>1</sup> To date, the mixed evidence for effects of prenatal smoking exposure on most outcomes in childhood and adulthood suggests that understanding shared familial context is a critical part of understanding health in childhood and adulthood. The finding regarding mortality benefits among low birthweight offspring of smokers has received and continues to receive substantial attention in the epidemiological literature,<sup>65,66</sup> and highlights that the role of family context, both shared and unshared, in producing bias continues to limit our ability to understand the effects of prenatal exposures. Family designs have yielded important insights, but they should be paired with longitudinal data sources that examine dynamic exposures within and across families and across pregnancies in order to fully understand the potential for downstream effects of the earliest exposures. Perhaps most important, it is critical to augment observational studies that rely on statistical adjustment to achieve comparability with observational studies that use strategic comparisons, instrumental variables and within-family estimation as alternative approaches to evaluate causal effects.<sup>21</sup>

## References

1. Yerushalmy J. The relationship of parents' cigarette smoking to outcome of pregnancy – implications as to the problem of inferring causation from observed associations. *Am J Epidemiol* 1971;93:443–56.
2. Yerushalmy J, Hilleboe HE. Fat in the diet and mortality from heart disease. *N Y State J Med* 1957;57:2343–54.
3. Yerushalmy J, Palmer CE. On the methodology of investigations of etiologic factors in chronic diseases. *J Chronic Dis* 1959;10:27–40.
4. Yerushalmy J. Statistical considerations and evaluation of epidemiological evidence. In: James G, Rosenthal T (eds). *Tobacco and Health*. Springfield, IL: Charles C. Thomas, 1962.
5. Susser E, Buka S, Schaefer CA *et al.* The early determinants of adult health study. *J Dev Orig Health Dis* 2011;2:311–21.
6. Broman SH. The collaborative perinatal project: an overview. In: Mednick SA, Harway M, Finello KM (eds). *Handbook of Longitudinal Research*. Vol 1. New York: Praeger Publishers, 1984.
7. van den Berg BJ, Christianson RE, Oechli FW. The California Child Health and Development Studies of the School of Public Health, University of California at Berkeley. *Paediatr Perinat Epidemiol* 1988;2:265–82.
8. Eskenazi B, Bergmann JJ. Passive and active maternal smoking during pregnancy, as measured by serum cotinine, and postnatal smoke exposure. I. Effects on physical growth at age 5 years. *Am J Epidemiol* 1995;142(Suppl 9):S10–18.
9. Eskenazi B, Trupin LS. Passive and active maternal smoking during pregnancy, as measured by serum cotinine, and postnatal smoke exposure. II. Effects on neurodevelopment at age 5 years. *Am J Epidemiol* 1995;142(Suppl 9):S19–29.

10. Keyes KM, Davey Smith G, Susser E. Associations of prenatal maternal smoking with offspring hyperactivity: causal or confounded? *Psychol Med* 2014;44:857–67.
11. Talati A, Bao Y, Kaufman J, Shen L, Schaefer CA, Brown AS. Maternal smoking during pregnancy and bipolar disorder in offspring. *Am J Psychiatry* 2013;170:1178–85.
12. Terry MB, Schaefer CA, Flom JD *et al.* Prenatal smoke exposure and mammographic density in mid-life. *J Dev Orig Health Dis* 2011;2:340–52.
13. DiFranza JR, Aligne CA, Weitzman M. Prenatal and postnatal environmental tobacco smoke exposure and children's health. *Pediatrics* 2004;113(Suppl 4):1007–15.
14. Kramer MS, Olivier M, McLean FH, Dougherty GE, Willis DM, Usher RH. Determinants of fetal growth and body proportionality. *Pediatrics* 1990;86:18–26.
15. Mitchell EA, Milerad J. Smoking and the sudden infant death syndrome. *Rev Environ Health* 2006;21:81–103.
16. Blackburn H, Labarthe D. Stories from the evolution of guidelines for causal inference in epidemiologic associations: 1953–1965. *Am J Epidemiol* 2012;176:1071–77.
17. Lilienfeld DE. Abe and Yak: the interactions of Abraham M. Lilienfeld and Jacob Yerushalmy in the development of modern epidemiology (1945–1973). *Epidemiology* 2007;18:507–14; discussion 515–16.
18. Donovan SJ, Susser E. Commentary: Advent of sibling designs. *Int J Epidemiol* 2011;40:345–49.
19. Susser M. *Causal Thinking in the Health Sciences. Concepts and Strategies in Epidemiology*. New York: Oxford University Press, 1973.
20. Cornfield J, Haenszel W, Hammond EC, Lilienfeld AM, Shimkin MB, Wynder EL. Smoking and lung cancer: recent evidence and a discussion of some questions. *J Natl Cancer Inst* 1959;22:173–203. Reprinted *Int J Epidemiol* 2009;38:1175–91.
21. Davey Smith G. Assessing intrauterine influences on offspring health outcomes: can epidemiological studies yield robust findings? *Basic Clin Pharmacol Toxicol* 2008;102:245–56.
22. Sullivan WC. A note on the influence of maternal inebriety on the offspring. 1899. Reprinted: *Int J Epidemiol* 2011;40:278–82.
23. Agrawal A, Scherrer JF, Grant JD *et al.* The effects of maternal smoking during pregnancy on offspring outcomes. *Prev Med* 2010;50:13–18.
24. Laucht M, Hohm E, Esser G, Schmidt MH, Becker K. Association between ADHD and smoking in adolescence: shared genetic, environmental and psychopathological factors. *J Neural Transm* 2007;114:1097–104.
25. 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.
26. Breslau N, Paneth N, Lucia VC, Paneth-Pollak R. Maternal smoking during pregnancy and offspring IQ. *Int J Epidemiol* 2005;34:1047–53.
27. Susser E, Eide MG, Begg M. Invited commentary: The use of sibling studies to detect familial confounding. *Am J Epidemiol* 2010;172:537–39.
28. Yang S, Lynch J, Susser ES, Lawlor DA. Birthweight and cognitive ability in childhood among siblings and nonsiblings. *Pediatrics* 2008;122:e350–58.
29. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;32:1–22.
30. Lewis SJ, Araya R, Davey Smith G *et al.* Smoking is associated with, but does not cause, depressed mood in pregnancy - a mendelian randomization study. *PLoS One* 2011;6:e21689.
31. Sexton M, Hebel JR. A clinical trial of change in maternal smoking and its effect on birthweight. *JAMA* 1984;251:911–15.
32. Permutt T, Hebel JR. Simultaneous-equation estimation in a clinical trial of the effect of smoking on birthweight. *Biometrics* 1989;45:619–22.
33. Hamilton BH. Estimating treatment effects in randomized clinical trials with non-compliance: the impact of maternal smoking on birthweight. *Health Econ* 2001;10:399–410.
34. Evans WN, Ringel JS. Can higher cigarette taxes improve birth outcomes? *J Public Econ* 1999;72:135–54.
35. Tyrrell J, Huikari V, Christie JT *et al.* Genetic variation in the 15q25 nicotinic acetylcholine receptor gene cluster (CHRNA5-CHRNA3-CHRNA4) interacts with maternal self-reported smoking status during pregnancy to influence birthweight. *Hum Mol Genet* 2012;21:5344–58.
36. Davey Smith G. Negative control exposures in epidemiologic studies. *Epidemiology* 2012;23:350–51.
37. Denson R, Nanson JL, McWatters MA. Hyperkinesia and maternal smoking. *Can Psychiatr Assoc J* 1975;20:183–87.
38. Yerushalmy J. Mother's cigarette smoking and survival of infant. *Am J Obstet Gynecol* 1964;88:505–18.
39. Yerushalmy J. Self-selection – a major problem in observational studies. In: Le Cam L, Neyman J, Scott EL (eds). *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 4: Biology and Health*. Berkeley, CA: University of California Press, 1972.
40. Macmahon B, Alpert M, Salber EJ. Infant weight and parental smoking habits. *Am J Epidemiol* 1966;82:247–61.
41. D'Onofrio BM, Van Hulle CA, Waldman ID *et al.* Smoking during pregnancy and offspring externalizing problems: an exploration of genetic and environmental confounds. *Dev Psychopathol* 2008;20:139–64.
42. 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.
43. 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.
44. 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.
45. Hogberg L, Cnattingius S, Lundholm C, D'Onofrio BM, Langstrom N, Iliadou AN. Effects of maternal smoking during pregnancy on offspring blood pressure in late adolescence. *J Hypertens* 2012;30:693–99.
46. Frisell T, Oberg S, Kuja-Halkola R, Sjolander A. Sibling comparison designs: bias from non-shared confounders and measurement error. *Epidemiology* 2012;23:713–20.
47. Keyes KM, Davey Smith G, Susser E. On sibling designs. *Epidemiology* 2013;24:473–74.

48. Juarez SP, Merlo J. Revisiting the effect of maternal smoking during pregnancy on offspring birthweight: a quasi-experimental sibling analysis in Sweden. *PLoS One* 2013;8:e61734.
49. Soderstrom L, Perez-Vicente R, Juarez S, Merlo J. Questioning the causal link between maternal smoking during pregnancy and offspring use of psychotropic medication: a sibling design analysis. *PLoS One* 2013;8:e63420.
50. Plomin R. Why are children in the same family so different from one another? *Behav Brain Sci* 1987;10:1–16.
51. Davey Smith G. Epidemiology, epigenetics and the ‘Gloomy Prospect’: embracing randomness in population health research and practice. *Int J Epidemiol* 2011;40:537–62.
52. Ebrahim S. Improving causal inference. *Int J Epidemiol* 2013;42:363–66.
53. Gaysina D, Fergusson DM, Leve LD *et al*. Maternal smoking during pregnancy and offspring conduct problems: evidence from 3 independent genetically sensitive research designs. *JAMA Psychiatry*. 2013;70:956–63.
54. Rutter M, Moffitt TE, Caspi A. Gene-environment interplay and psychopathology: multiple varieties but real effects. *J Child Psychol Psychiatry* 2006;47:226–61.
55. Kristensen P, Bjerkedal T. Explaining the relation between birth order and intelligence. *Science* 2007;316:1717.
56. Yerushalmy J. Infants with low birthweight born before their mothers started to smoke cigarettes. *Am J Obstet Gynecol* 1972;112:277–84.
57. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008;168:656–64.
58. Cole SR, Platt RW, Schisterman EF *et al*. Illustrating bias due to conditioning on a collider. *Int J Epidemiol* 2010;39:417–20.
59. Plomin R. Commentary: Why are children in the same family so different? Non-shared environment three decades later. *Int J Epidemiol* 2011;40:582–92.
60. Plomin R, Daniels D. Why are children in the same family so different from one another? *Int J Epidemiol* 2011;40:563–82.
61. Conley D. Commentary: Reading Plomin and Daniels in the post-genomic age. *Int J Epidemiol* 2011;40:596–98.
62. Turkheimer E. Commentary: variation and causation in the environment and genome. *Int J Epidemiol* 2011;40:598–601.
63. Sesardic N. Commentary: An explosion without a bang. *Int J Epidemiol* 2011;40:592–96.
64. Skjaerven R, Wilcox AJ, Russell D. Birthweight and perinatal mortality of second births conditional on weight of the first. *Int J Epidemiol* 1988;17:830–38.
65. Basso O, Wilcox AJ. Intersecting birthweight-specific mortality curves: solving the riddle. *Am J Epidemiol* 2009;169:787–97.
66. Hernandez-Diaz S, Wilcox AJ, Schisterman EF, Hernan MA. From causal diagrams to birthweight-specific curves of infant mortality. *Eur J Epidemiol* 2008;23:163–66.