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

Katherine M Keyes,1,2* George Davey Smith3 and Ezra Susser1,2

1Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, USA, 2New York State Psychiatric Institute, New York, USA and 3MRC 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

The seminal 1971 paper by Jacob Yerushalmy,1 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.2–4 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). 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. 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, hyperactivity in adolescence, bipolar disorder and lower mammographic density in adulthood. 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, but with few exceptions (e.g. low birthweight, sudden infant death syndrome), 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. 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, 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.

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. 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. 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). The reasons to use a family-based strategy are 2-fold. First, there may be sources of genetic
non-comparability across families. For example, there is
evidence for shared genetic variance influencing both
smoking and hyperactivity.23–25 Thus, women who smoke
in pregnancy may also carry a higher risk of hyperactivity
themselves, and may pass on genetic risk factors for hyper-
activity to offspring. Second, other sources of non-compar-
ability that are shared within a family, such as the micro-
family rearing environment, and subtleties of social pos-
tion, are difficult to measure and fully control in standard
analytical approaches. In some contexts, control for mater-
nal cognition and education seem to be adequate to adjust
for family-level factors,26–28 but it is impossible to know
whether such adjustments are fully sufficient to eliminate
non-comparability. Although we still do not have defini-
tive strategies for untangling the underlying causal rela-
tionships, at least within a single observational study,
various approaches taken together are often complemen-
tary, as Yerushalmy suggested. When statistically-adjusted
results triangulate with those estimated from other designs
(e.g. Mendelian randomization29,30 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,31–33 non-genetic instrumen-
tal variables analyses,34 Mendelian randomization stud-
ies35 and maternal-paternal comparisons.21

Yerushalmy was among the first that we can find in the
literature to use paternal smoking status as a ‘negative con-
trol’, capturing aspects of the family context to study off-
spring health,21,36–38 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).4 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 off-
spring 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 asso-
ciation of maternal smoking with offspring hyperactivity
than for paternal smoking would be observed.

In Figures 1 and 2 we reprint the 1962 publica-
tion findings, as they demonstrate the clarity of the
maternal-paternal comparison approach and illustrate the
difficulty of ‘self-selection’, as Yerushalmy terms it,39 in in-
terpreting the results. Yerushalmy documents that the
probability of low birthweight (defined in his analysis as
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 birth-weight 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.4 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 paper40 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.10,21,36 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.18 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,41–43 cognition,42 body size44 and blood pressure.45 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.46,47 Indeed, discordant sibling studies confirm well-documented effects of prenatal smoking on lower offspring birthweight,48 but effect estimates are lower than have been documented in the literature.49 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;50,51 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.29,52 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,25 whereas another analysis reported that it was related to offspring conduct disorder symptoms at mean age 5 years.53 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.53 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.18,51 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),50,51,54 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,55 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,57,58 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.40 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.56 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, 1987,50 reprinted in IJE with commentaries in 2011).59-63 These stochastic elements can be adequately modeled.51

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.64 Similarly, in sibling comparisons, higher birthweight of a sibling is, on average, associated with higher and increase infant mortality.51 MacMahon and colleagues noted this potential explanation in their 1966 analysis of birthweight and mortality,65,66 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’.7 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,65,66 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.21

References


60. Plomin R, Daniels D. Why are children in the same family so different from one another? *Int J Epidemiol* 2011;40:563–82.