Why Evidence for the Fetal Origins of Adult Disease Might Be a Statistical Artifact: The “Reversal Paradox” for the Relation between Birth Weight and Blood Pressure in Later Life

Yu-Kang Tu¹,², Robert West¹, George T. H. Ellison³, and Mark S. Gilthorpe¹

¹ Biostatistics Unit, Centre for Epidemiology and Biostatistics, University of Leeds, Leeds, United Kingdom.
² Leeds Dental Institute, University of Leeds, Leeds, United Kingdom.
³ St. George’s Hospital Medical School, London, United Kingdom.

Received for publication March 11, 2004; accepted for publication June 3, 2004.

Inverse associations observed between low birth weight and markers of chronic disease in later life have generated what is termed the “fetal origins of adult disease hypothesis” (1, 2). The idea is that an unfavorable environment, or insults during fetal life, might induce lifetime effects on the subsequent development of body systems and hence give rise to major disease processes such as hypertension (2), diabetes (3), arteriosclerosis (4), asthma (5), and obesity (6). Over the last decade, many studies have been undertaken in many parts of the world to examine these proposed relations (7). Although some researchers have questioned the biologic basis of the hypothesis as well as its clinical importance, the concept that low birth weight is an independent risk factor for a range of chronic diseases in later life is now widely recognized as scientifically plausible and linked to poor fetal nutrition (8). One consequence of this seemingly plausible mechanism is that the fetal origins hypothesis is increasingly viewed as an important issue for public health and preventive medicine (9).

Nonetheless, two recent articles outlined substantive challenges to the fetal origins hypothesis (7, 8). One article raised concerns about the statistical methodology used and the improper interpretation of epidemiologic analyses invoked in support of the hypothesis (8). The second article suggested that the inverse association between birth weight and adult diseases might “chiefly reflect the impact of random error” in the measurement of birth weights, as well as “selective emphasis on particular results” (by which they meant publication bias in favor of analyses describing inverse relations between birth weight and adult blood pressure), and “inappropriate adjustment for current weight and for [other] confounding factors” (7, p. 659). A recent meta-analysis also
indicated that the relation between blood pressure and birth weight might suffer publication bias because small studies were more likely to report stronger inverse associations (10). Indeed, although a number of retrospective studies have found a direct relation between birth weight and adult health outcomes, others have found that a significant relation emerged only after adjusting for subsequent body size (notably current adult weight or body mass index) in the statistical analyses. It is the latter that forms the focus of this article.

The inappropriate use of statistical adjustment(s) for so-called confounders is an important source of potential bias among observational studies of the fetal origins of adult disease hypothesis. We sought to demonstrate how adjustment for current body mass (or any similar measure of current body size) is inappropriate because it is not a true “confounder” but part of the causal pathway between birth weight and adult blood pressure. The resulting phenomenon has been given the generic name the “reversal paradox” (or the “amalgamation paradox”) (11), although it is perhaps not widely known as either. Whatever the name, the reversal paradox makes it very challenging to correctly interpret the findings of observational studies of the causal links between fetal growth and adult disease where covariates are inappropriately treated as confounders.

Because the inverse association between birth weight and blood pressure is considered the most statistically consistent of the associations between birth weight and health in later life (7), our article adopts this example for illustration. Nevertheless, what follows is essentially applicable to all health outcomes in later life and to any other descriptive epidemiologic analyses in which similar statistical adjustments are undertaken inappropriately.

MATERIALS AND METHODS

To avoid complicating the issues surrounding adjustment for multiple confounders, such as age and gender, we decided to consider only a single hypothetical sample comprising adult males of equal age. Synthetic data for these men were generated for three variables—birth weight, current adult weight, and adult systolic blood pressure—and for three scenarios in which there is 1) no direct relation between birth weight and blood pressure (i.e., the Pearson correlation between birth weight and blood pressure is 0); 2) a modest inverse relation between birth weight and blood pressure (i.e., the Pearson correlation between birth weight and blood pressure is –0.05); and 3) a modest positive relation between birth weight and blood pressure (i.e., the Pearson correlation between birth weight and blood pressure is 0.05). Mean values for the three variables and their standard deviations were derived from the literature (12) and from the results of surveys conducted by the United Kingdom Department of Health (http://www.doh.gov.uk): birth weight = 3.38 kg (standard deviation, 0.57 kg), current weight = 82.60 kg (standard deviation, 14.75 kg), and systolic blood pressure = 130.0 mmHg (standard deviation, 11.2 mmHg).

Since the effects of the reversal paradox depend on the pairwise correlations of birth weight, current weight, and blood pressure, each scenario was simulated for a range of different assumptions. To illustrate the impact of the reversal paradox on changes in the relation between birth weight and blood pressure after statistically adjusting for current weight, we assumed three different values for the birth weight–current weight correlation (0.15, 0.25, and 0.35) and four different values for the current weight–blood pressure correlation (0.15, 0.25, 0.35, and 0.45). The range of correlation values adopted was motivated by typical values encountered in the literature. The theoretical basis of these simulations is outlined in the Appendix.

With adult blood pressure as the outcome variable, we estimated a median value for the partial regression coefficient with birth weight after adjusting for current adult weight, using simulations based on a sample of 500 persons—the mid-range sample size of previous empirical studies (7)—and with 1,000 iterations for each scenario. All simulations and statistical evaluations were performed by using the statistical package R, version 1.7.1 (13). For each scenario, the function “mvrnorm” in the MASS package in R was used to generate multivariate normal data using the mean values, standard deviations, and given correlation matrix for the three variables birth weight, current weight, and blood pressure (14).

RESULTS

Scenario 1: the correlation between blood pressure and birth weight is 0

When the Pearson correlation between blood pressure and birth weight was 0, the simple (bivariate) regression coefficient of birth weight, for blood pressure regressed on birth weight, was unsurprisingly close to 0. In contrast, when current weight was included in the model, the regression coefficient for birth weight became negative, and the magnitude of this coefficient increased as the birth weight–current weight and blood pressure–current weight correlations increased (table 1). When both of the latter correlations were set to 0.15, the estimated effect of a 1-kg increase in birth weight was a 0.42-mmHg reduction in adult blood pressure. With the birth weight–current weight correlation increased to 0.35 and the blood pressure–current weight correlation increased to 0.45, the estimated effect of a 1-kg increase in birth weight was a 3.51-mmHg reduction in adult blood pressure.

Scenario 2: the correlation between blood pressure and birth weight is –0.05

When a modest inverse relation between birth weight and blood pressure was adopted, the theoretical value of the simple (bivariate) regression coefficient of birth weight, for blood pressure regressed on birth weight, was –0.98 mmHg/kg. This value can be derived by multiplying the correlation coefficient by the ratio of standard deviation for blood pressure to that of the standard deviation for birth weight (–0.05 × [11.2 ÷ 0.57]), and the results of the simulations were very close to this (ranging from –0.95 to –1.01; table 2). When the birth weight–current weight and blood pressure—
current weight correlations were both 0.15, the estimated effect of a 1-kg increase in birth weight was a 1.48-mmHg reduction in adult blood pressure, similar to that estimated by a previous meta-analysis (7). With the birth weight–current weight correlation increased to 0.35 and the blood pressure–current weight correlation increased to 0.45, increasing birth weight by 1 kg yielded a 4.47-mmHg reduction in adult blood pressure.

Scenario 3: the correlation between blood pressure and birth weight is 0.05

When a modest positive relation between birth weight and blood pressure was adopted, the theoretical value of the coefficient of birth weight, for blood pressure regressed on birth weight, was 0.98 mmHg/kg (0.05 × [11.2 ÷ 0.57]), as anticipated from the symmetry of scenario 3 with scenario 2; refer to table 3. When the birth weight–current weight and blood pressure–current weight correlations were both 0.15, a 1-kg increase in birth weight was associated with a 0.55-mmHg increase in adult blood pressure. When the birth weight–current weight and blood pressure–current weight correlations were both 0.25, the regression coefficient was reversed, and increasing birth weight by 1 kg became associated with a 0.26-mmHg reduction in adult blood pressure. With the birth weight–current weight and blood pressure–current weight correlations increased to 0.35 and 0.45, respectively, the estimated effect of a 1-kg increase in birth weight was a 2.41-mmHg reduction in adult blood pressure.

**DISCUSSION**

For categorical variables within the field of probability and statistical science, the reversal paradox is best known as “Yule’s paradox” or “Simpson’s paradox.” George U. Yule noticed this phenomenon as early as 1903 (15), when he referred to a paper published by Karl Pearson et al. dating from 1899 (16). The issue was later mentioned in a 1951 paper by Edward H. Simpson on the way in which the relation between two variables changed after a third variable was factored into a two-by-two contingency table (17). When such data are analyzed by regression methods, the reversal paradox is more often referred to as “Lord’s paradox,” particularly within the behavioral sciences (18), ever since Frederic M. Lord published his 1967 paper on this phenomenon with respect to the use of analysis of covariance (19).

Within any generalized linear modeling framework, this phenomenon is more generally known in the statistical literature as the “suppression effect,” with the third variable termed a “suppressor” (20, 21). Thus, in whatever form and under whatever name, the reversal paradox has been recognized ever since the statistical methods of correlation and regression became established. Indeed, the paradox was discussed in 1910 by Karl Pearson and Arthur C. Pigou, Professor of Political Economics at Cambridge University in the United Kingdom, when they debated the role of parental alcoholism and its impact on the performance of children (11). However labeled, the paradox has been extensively explored in the statistical literature, especially in the behav-

**TABLE 1.** Association between birth weight and adult blood pressure in simple regression models and multiple regression models after adjustment for adult body weight in scenario 1†

<table>
<thead>
<tr>
<th>Simple regression coefficient (mmHg/kg)</th>
<th>Birth weight–current weight correlation</th>
<th>Multiple regression coefficient (mmHg/kg): blood pressure–current weight correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median 95% CI†‡</td>
<td>Median 95% CI</td>
<td>Median 95% CI</td>
</tr>
<tr>
<td>0.02†‡</td>
<td>1.69, 1.77</td>
<td>0.15</td>
</tr>
<tr>
<td>0.07†‡</td>
<td>1.79, 1.74</td>
<td>0.25</td>
</tr>
<tr>
<td>0.01†‡</td>
<td>1.76, 1.82</td>
<td>0.35</td>
</tr>
</tbody>
</table>

† The bivariate correlation between birth weight and blood pressure is 0.
‡ CI, confidence interval.

**TABLE 2.** Association between birth weight and adult blood pressure in simple regression models and multiple regression models after adjustment for adult body weight in scenario 2†

<table>
<thead>
<tr>
<th>Simple regression coefficient (mmHg/kg)</th>
<th>Birth weight–current weight correlation</th>
<th>Multiple regression coefficient (mmHg/kg): blood pressure–current weight correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median 95% CI†‡</td>
<td>Median 95% CI</td>
<td>Median 95% CI</td>
</tr>
<tr>
<td>–1.01†‡</td>
<td>–2.75, 0.70</td>
<td>0.15</td>
</tr>
<tr>
<td>–0.95†‡</td>
<td>–2.87, 0.69</td>
<td>0.25</td>
</tr>
<tr>
<td>–0.96†‡</td>
<td>–2.58, 0.66</td>
<td>0.35</td>
</tr>
</tbody>
</table>

† The bivariate correlation between birth weight and blood pressure is –0.05.
‡ CI, confidence interval.

‡ The expected regression coefficient should be –0.98 (–0.05 × [11.2 ÷ 0.57]).
TABLE 3. Association between birth weight and adult blood pressure in simple regression models and multiple regression models after adjustment for adult body weight in scenario 3

<table>
<thead>
<tr>
<th>Simple regression coefficient (mmHg/kg)</th>
<th>Birth weight–current weight correlation</th>
<th>Multiple regression coefficient (mmHg/kg): blood pressure–current weight correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median 95% CI†</td>
<td>0.15</td>
<td>Medain 95% CI</td>
</tr>
<tr>
<td>0.97†</td>
<td>0.15</td>
<td>0.24 –1.43, 1.84</td>
</tr>
<tr>
<td>1.00†</td>
<td>0.25</td>
<td>–0.26 –2.02, 1.38</td>
</tr>
<tr>
<td>0.95†</td>
<td>0.35 –0.06 –1.85, 1.69</td>
<td>–0.85 –2.65, 0.90</td>
</tr>
</tbody>
</table>

† The expected regression coefficient should be 0.98 (0.05 × [11.2 + 0.57]).

ioral sciences (22, 23), yet comparatively few of these analyses acknowledge that they are in fact different manifestations of the same phenomenon; that is, they are all just one paradox.

Moreover, while the original definition and naming of the reversal paradox drew on the notion that the direction of any relation between two variables is reversed after a third variable is introduced, it may nevertheless be generalized to scenarios in which the relation between two such variables is enhanced, not reversed nor reduced, after a third variable is introduced. This difference is not widely appreciated, which is perhaps why the impact of the reversal paradox is not always recognized in empirical, rather than theoretical, statistical analyses. However, the potential problem of statistical adjustment in the fetal origins hypothesis has been known for some time. Paneth et al. (24, 25) specifically addressed the issue of “overcontrolling” for current body mass index when analyzing the relation between birth weight and disease risk in later life. To illustrate the reversal paradox and explore this point further, we used simulations of hypothetical data to mimic three scenarios pertaining to the fetal origins hypothesis for adult blood pressure. For those interested in the statistical theory underpinning these ideas, refer to the Appendix.

The reversal paradox invokes bias due to the inappropriate “controlling” of alleged confounders that are not in fact “true confounders” (26). The concept of what constitutes a confounder has been revised in recent years, with greater emphasis given to the definition of “causality” in the associations among outcomes, exposures, and confounders. Detailed expositions on this issue have emerged only recently (27–29), and these stricter, revised definitions of what constitutes a confounder may not have been disseminated, or universally accepted, throughout the discipline of epidemiology. Consequently, the reason why the reversal paradox is a problem in some instances and not others may be more of a philosophical issue than a statistical one. Indeed, the principal issue of statistical adjustment pertinent to the fetal origins hypothesis is the one surrounding the causal pathway and the position within it of current adult body weight as an alleged confounder. If one defines a causal pathway as the chain of events or factors leading in sequence to an outcome, it only makes sense to examine an outcome in relation to any one point along the causal pathway (27, 30).

We differentiated between two different, but complementary causal pathways: 1) low birth weight affecting blood pressure directly (e.g., poor nutrition in utero having an irreversible impact on the subsequent development of cardiovascular systems); and 2) low birth weight affecting blood pressure via birth weight’s impact on current weight (e.g., through a genetic link between size at birth and current adult body size), which in turn is causally related to high blood pressure (i.e., birth weight → current weight → blood pressure). In the latter model, it is sensible to examine either the relation between birth weight and adult blood pressure or the relation between current weight and adult blood pressure, in isolation. It is not sensible to examine the relation between birth weight and adult blood pressure while controlling for current weight, because adult weight lies on the causal pathway between the outcome (blood pressure) and the exposure (birth weight). To statistically “adjust” for current weight while exploring the impact of birth weight on adult blood pressure invokes the reversal paradox.

In a slightly different formulation of the fetal origins hypothesis, some researchers have argued that persons of a relatively low birth weight and a relatively high adult body size (be it weight, height, or body mass index) ought to be considered a “high-risk” group for cardiovascular disease in later life (31). Indeed, this is another possible interpretation of the statistical models commonly used in analyses exploring the fetal origins hypothesis. However, as discussed previously, it is not appropriate to include current weight as a confounder in these analyses if the relation between birth weight and adult blood pressure is the primary interest because weight gain, like current adult weight, is in the causal pathway from birth weight to adult blood pressure. Nonetheless, the same statistical model is useful and appropriate if the relation between adult blood pressure and current weight is the primary interest; the precision of estimating the association between adult blood pressure and current weight can be enhanced by including birth weight in the analyses. This asymmetric utility of adjusting for “confounding” occurs because the adult body weight of some adults is likely to be greater because they were born heavier. For these adults, their greater body weight may not be associated with the factors that give rise to higher adult

blood pressure but with their greater size at birth (which might, for example, be genetically determined). Under these circumstances, the association between adult blood pressure and current weight will be attenuated, and adjusting for birth weight deals with the way in which the relation between adult blood pressure and current weight would otherwise be diluted. Nevertheless, it should be stressed that, in this analytic model, the relation between adult blood pressure and birth weight does not have any empirical utility.

In the epidemiologic literature, body size measurements (be they body weight, body height, or body mass index) are frequently considered confounders for health-related outcomes, presumably because there are well-known allometric relations between body size and function (32). However, there appears to be no consistent practice in the literature as to when and how such variables should (or should not) be “controlled for,” and most studies do not offer any justification for their choice of confounders. For instance, a recent study of the fetal origins hypothesis that adjusted for current body weight and body mass index simultaneously found a significant negative relation between birth weight and adult systolic blood pressure, even though there was no significant (simple, bivariate) correlation between them prior to adjustment (33). Furthermore, because body weight is the numerator of body mass index, including both variables in the analyses simultaneously means that they will have undergone mathematical coupling (34, 35). A recently published meta-analysis, along with other studies, shows that removing the adjustment for current weight (or a comparable measure of current body size) reduces the association between birth weight and adult blood pressure (7); in some instances, the association is no longer statistically significant (36). Our scenario 2 might therefore be closest to reality.

Note that the estimated effect size in this simulated study, or in any empirical study, is affected by the sample ratio of the blood-pressure standard deviation to that of the birth-weight standard deviation. If this ratio is large in any particular study, for instance, when the adult age range is wide and thereby yields a wider range of adult blood pressures, the effect size (i.e., the extent of bias) caused by the reversal paradox will be exaggerated. On the other hand, if the sample ratio of blood-pressure standard deviation to birth-weight standard deviation is small, as, for instance, among studies of children or young adults in which blood-pressure variation tends to be smaller than among older adults, the effect size caused by the reversal paradox will be diminished. Nevertheless, testing the significance of an association is affected by sample size, and it is well known that even a small effect size will be statistically significant if the sample size is large enough. Most studies examining the fetal origins of adult blood pressure have sufficiently large samples to yield statistical significance for relatively small effect sizes. Thus, the exaggerating effect of the reversal paradox tends to give the misleading impression that the relation between blood pressure and birth weight, after adjustment for current weight, is not only statistically significant (because of the power available to detect the biased difference from 0) but also biologically and clinically significant (as a result of the biased effect size caused or enhanced by the reversal paradox).

Frequently, statistical methodology in studies of the fetal origins hypothesis is at risk of the reversal paradox, thereby bringing into question the validity and reliability of some of the purported evidence. It is thus difficult, if not impossible, to compare results across studies where so many varied attempts have been made to control for confounders, without consistent reasoning concerning the choice of these confounders. This difficulty does not invalidate the fetal origins hypothesis per se; rather, it implies that any direct interpretation of inverse relations between birth weight and any adult condition, while adjusting for current adult body measurements, cannot be taken to mean that birth weight has a direct impact on the adult outcome. For our hypothetical example, a more appropriate interpretation is that, if all babies grew to the same size in adulthood, lower-birth-weight babies would, on average, have higher blood pressure in adulthood. Yet, this conclusion is counterfactual since low-birth-weight babies will, on average, be smaller than heavy-birth-weight babies in adulthood; there is a positive correlation between birth weight and adult weight, as studies supporting the fetal origins hypothesis have shown.

As an understated and poorly recognized issue, the reversal paradox, in whatever form it takes, has the potential to severely affect data analyses undertaken in empirical research, which increasingly rely on the methods of generalized linear modeling of observational (i.e., nonrandomized) data. Our simulations highlight this issue with respect to the fetal origins of adult disease hypothesis, where the paradox is perhaps instrumental in generating a great deal of the evidence cited within the hypothesis’ burgeoning orthodoxy. However, the aim of this article was not to refute the fetal origins hypothesis but to remind epidemiologists that, to arrive at the correct interpretation of evidence supporting or refuting this hypothesis, one must fully understand the statistical methods used and the implications of making statistical adjustment for “confounders.”

References

APPENDIX

Denoting adult blood pressure by BP, birth weight by BW, and current weight by CW, and assuming that the three variables follow a trivariate normal distribution, we may write BW = a_1 + b_1 CW + e_1 and BP = a_2 + b_2 CW + e_2. The unconditional covariance between BW and BP is thus b_1b_2 var(CW) + cov(e_1, e_2). If both BW and BP are positively correlated with CW, then b_1 and b_2 must both be positive, and hence so is their product. Thus, the only way to obtain a zero covariance between BW and BP—that is, a zero bivariate correlation between them, as was imposed in scenario 1—is for cov(e_1, e_2) to be negative. Now, if we look at cov(BW, BP) while conditioning on CW (as in a regression model that includes both CW and BW as predictors), the conditional covariance is equal to cov(e_1, e_2), which we have just indicated must be negative. The sign of the coefficient of BW as determined by cov(BW, BP) must therefore be negative in the model that includes CW, provided that both BW and BP are positively correlated with CW and that BW and BP are also uncorrelated.

If we denote growth weight by GW, as per Lucas et al. (8), the following models could be rearranged:

\[ \text{BP} = \beta_0 + \beta_1 \text{BW} + \beta_2 \text{CW}, \]  

and \[ \text{BP} = \beta_0 + \beta_1 \text{BW} + \beta_2 (\text{BW} + \text{GW}) = \beta_0 + (\beta_1 + \beta_2) \text{BW} + \beta_2 \text{GW}. \]  

In model 2, the regression coefficient for birth weight is attenuated because \( \beta_2 \) is always positive. In other words, whenever weight gain after birth (i.e., growth weight) is adjusted for, the inverse association between birth weight and blood pressure will be less than that obtained while “controlling” for current weight. However, the two approaches—controlling for current weight or growth weight—are no different from a statistical viewpoint since birth weight is contained in some aspect within current adult weight. The difficult question then arises: which coefficient is the more accurate estimate of the contribution of birth weight to adult blood pressure?