Invited Commentary

Invited Commentary: Coming Out of the Box

Pam Factor-Litvak and Arnold Sher

Initially submitted February 4, 2009; accepted for publication February 10, 2009.

The authors apply the analogy of a “black box” from systems theory to epidemiologic problems. They highlight this process using the example of associations between sequential measures of body size and systolic blood pressure. Several other examples of the use of structural equation modeling or path analysis are described. Finally, the authors highlight several requirements for using these methods to “come out of the box.”

birth weight; blood pressure; body mass index; child; epidemiologic methods; growth

The paradigm of “black box” epidemiology, hotly debated in the last 2 decades (1), is slowly giving way to a more inclusive approach encompassing explanations on the macro, micro, and life course levels—so-called “eco-epidemiology” (2). While results based on black box paradigms continue to be important for public health interventions (1), investigators are increasingly concerned with distal preventive strategies. In this spirit, the accompanying paper by Gamborg et al. (3) illustrates 1 possible approach to the analysis and presentation of life course data.

In systems theory, the black box is intermediate between the inputs (or risk factors) and the output (or disease outcome). Transformations of the inputs occur within the black box, resulting in the output; one could liken these transformations to biologic or social mechanisms. Consider a simple system with 1 input, leading to the black box, and 1 output. In this scenario, there are 2 basic approaches to changing the output. First, changing the input will change the output. Take, for example, the early epidemiologic studies of smoking and lung cancer (4, 5). In the absence of detailed biologic evidence relating cigarette smoke components to lung cancer, results indicated that promoting smoking cessation programs would reduce the incidence of lung cancer. Second, changing the processing within the box will also change the outcome. This requires knowledge of the box structure, its workings, its pathways, and its processes. It is this latter mode of change that Gamborg et al. (3) address.

In their example, Gamborg et al. (3) address the hypothesis that the observed associations between birth weight and adult blood pressure are mediated by body size during the life course. This hypothesis is based on a multitude of observations suggesting that birth weight is inversely associated with adult adverse outcomes (6–8) and that higher body mass index (weight (kg)/height (m)²) is also associated with adult adverse outcomes (9–11). Gamborg et al. (3) and others (12, 13) posit that the association between birth weight and, for example, systolic blood pressure in adulthood may be partly spurious because of misspecification of the regression models. They propose the use of life course path analysis to describe the process behind the association.

The box that Gamborg et al. (3) open describes the mechanism by which birth weight leads to adult body mass index. Using a rich set of longitudinal data, this particular mechanism, most likely one of many relating birth weight to health outcomes, describes the successive relations between body mass at various points in childhood and adolescence and adult systolic blood pressure. Thus, as indicated by the authors, the estimated effects of each measure of body mass index on adult systolic blood pressure include those mediated by future body mass index. The results, as shown in Gamborg et al.’s path analysis life course plot (Figure 3, part B, in their paper (3)), strongly suggest that the association between birth weight and adult systolic blood pressure is small compared with that of future body size. Further, the results suggest that the associations between body mass index and adult systolic blood pressure become stronger around the time of puberty. These results are particularly important in targeting weight reduction and physical activity interventions for children and teenagers.

This is not, however, the first use of path analysis in epidemiology. In their 1975 study of the Dutch Hunger

Correspondence to Dr. Pam Factor-Litvak, Department of Epidemiology, Mailman School of Public Health, Columbia University, 722 West 168 Street, Room 1614, New York, NY 10032 (e-mail: prf1@columbia.edu).
Winter of 1944–1945, Stein et al. (14) used path analysis to explain the relations between maternal prepregnancy weight, placental weight, birth weight, birth length, and head circumference. Those results suggested that in the context of inadequate maternal nutrition, the first impact was on maternal weight, and only after a lower threshold of maternal weight was reached was the weight of the conceptus reduced. Thus, Stein et al. began to open the box relating prenatal nutrition to fetal growth.

More recent examples include investigations of the life course pathways to literacy and numeracy problems (15). Here, the black box concerned the associations between social class of origin (i.e., father’s occupation) and literacy/numeracy problems at midlife in 2 British birth cohorts (the 1946 and 1958 cohorts). The a priori path included direct effects of cognitive ability in childhood, educational attainment, and occupation on literacy and numeracy; antecedent to all was father’s occupation. Direct pathways were also hypothesized from father’s occupation to cognitive ability, educational attainment, and occupation, but not to literacy/numeracy problems. Data analysis confirmed these hypothesized pathways in both cohorts, although the magnitudes of the associations differed because of a contextual expansion of secondary education in the latter.

In another example, Christ et al. (16) examined 2 alternative pathways from self-reported visual impairment to mortality—the first pathway operating via self-rated health and the second via disability. Using data from the National Health Interview Survey, linked to the National Death Index, the investigators found that although visual impairment is directly associated with mortality, the association becomes stronger once the indirect effects through self-reported health and disability are taken into account.

A final example concerns the estimation of direct and indirect effects of neighborhood characteristics on drug use and high-risk sexual behaviors (17). In this a priori hypothesized model, neighborhood characteristics such as reported vandalism, vacant housing, and robberies were hypothesized to have both direct relations with drug use and high-risk sexual behavior and an indirect association with high-risk sexual behavior operating via psychological distress and drug use. Results from this study indicated that neighborhood interventions as well as individual interventions are necessary to address these public health issues.

These examples, along with the study by Gamborg et al. (3), highlight several features of the use of path analysis to open the black box. First, paths of interest are stated in an a priori fashion. This requires knowledge of the literature and collaboration from many different perspectives—social, psychological, and biological—to delineate a plausible model. Such modeling is in the spirit of eco-epidemiology. Second, several path models can be constructed to test alternative models. Further, it is likely that these models will share components and be interlinked. It is also likely that each model will account for both shared and unique variance in the outcome. Finally, because social, psychological, and biological processes are complex, it is also likely that the model(s) will not account for 100% of the variance in the outcome. Residual variance may be attributed to both measurement error and mechanisms hidden at the bottom of the box.

The current focus on life course epidemiology and the social determinants of health demands a rethinking of data analytic strategies. Such strategies include but are not limited to structural equation modeling, path analysis, and multilevel modeling. Papers such as that of Gamborg et al. (3) remind us of alternative methods for “coming out of the box.”

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY (Pam Factor-Litvak); and ICSD Consulting, Millburn, New Jersey (Arnold Sher).

Conflict of interest: none declared.

REFERENCES

13. Tu YK, West R, Ellison GT, et al. Why evidence for the fetal origins of adult disease might be a statistical artifact: the “reversal paradox” for the relation between birth weight and


