

Letters to the Editor

RE: "NEED FOR MORE INDIVIDUAL-LEVEL META-ANALYSES IN SOCIAL EPIDEMIOLOGY: EXAMPLE OF JOB STRAIN AND CORONARY HEART DISEASE"

In their article on meta-analyses in social epidemiology, Kivimäki and Kawachi (1) make a strong case for systematically pooling data from observational studies in the social epidemiology field, and Kivimäki et al.'s demonstration of publication bias in the literature on job strain and coronary heart disease risk is compelling (2). However, we take issue with the analogy between individual-level meta-analyses in social epidemiology and the success of meta-analyses in the field of genetic epidemiology. As Kivimäki and Kawachi point out, the profusion of nonreplicable candidate gene associations (3) has now been replaced by the identification of reliable and definite associations from large collaborative studies. The situation for studies of germ-line genetic variation is based upon the fact that, with basic precautions being taken to avoid population stratification, there is little confounding (4), no reverse causation (the early stages of a disease process influencing the apparent exposure), and lower risk of several other biases that plague observational epidemiology (5). In genetic epidemiology, therefore, low statistical power and the selective publication of false-positive associations are the major concerns.

The situation is fundamentally different in meta-analyses of nongenetic epidemiology, where confounding and bias are the major issue with respect to a causal interpretation of associations (6). Here the pooling of data from several observational studies may hamper appropriate confounder adjustment, due to the need to work with a lowest common denominator after "harmonizing" the data across studies. This is illustrated by the meta-analysis of studies of job strain and coronary heart disease (2) that motivated Kivimäki and Kawachi's commentary (1). The assessment of socioeconomic status was based on occupational titles (in 1 study on education), and socioeconomic status was categorized into low, intermediate, and high. Tobacco smoking and alcohol consumption were similarly crudely categorized, despite the fact that many of the individual studies had collected detailed information on these variables. Therefore, it is not surprising that many precisely estimated effects from meta-analyses of observational studies are of no causal relevance (6). Investigation of such associations requires different techniques, such as studies in settings where confounding structures differ (7), the use of negative control outcomes or exposures (8–10), Mendelian randomization (11), and other instrumental variable approaches. In this regard, it is important to note that non-germ-line genomic measures, such as epigenetic profiles, suffer from the usual limitations of observational epidemiology (12, 13) and would also be served badly by simply pooling data across different studies.

The use of large-scale individual participant data meta-analyses of observational studies was pioneered by the Prospective Studies Collaboration (14) and the Emerging Risk Factors Collaboration (15). These initiatives predated by a decade the recognition of the value of pooling to increase power to detect small effect sizes in genetic epidemiology. In areas where randomized trial data have already confirmed that the "exposure"–outcome association is causal, observational data can be pooled to examine the range and generalizability of the causal association, as is the case with high blood pressure and cardiovascular disease. In areas where no randomized trials exist—for example, C-reactive protein and cardiovascular disease—the pooling of large volumes of data to create very precise estimates with very small *P* values has not proven helpful in making causal inferences (15). Indeed, a separate C-reactive protein coronary disease genetics consortium was established to test the C-reactive protein hypothesis using a Mendelian randomization design (16).

ACKNOWLEDGMENTS

Conflict of interest: none declared.

REFERENCES

1. Kivimäki M, Kawachi I. Need for more individual-level meta-analyses in social epidemiology: example of job strain and coronary heart disease. *Am J Epidemiol.* 2012;177(1):1–2.
2. Kivimäki M, Nyberg ST, Batty GD, et al. Job strain as a risk factor for coronary heart disease: a collaborative meta-analysis of individual participant data. *Lancet.* 2012;380(9852):1491–1497.
3. Colhoun HM, McKeigue PM, Davey Smith G. Problems of reporting genetic associations with complex outcomes. *Lancet.* 2003;361(9360):865–872.
4. Davey Smith G, Lawlor DA, Harbord R, et al. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Med.* 2008;4(12):1985–1992.
5. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol.* 2003;32(1):1–22.
6. Egger M, Schneider M, Davey Smith G. Spurious precision? Meta-analysis of observational studies. *BMJ.* 1998;316:140–144.
7. Brion MJ, Lawlor DA, Matijasevich A, et al. What are the causal effects of breastfeeding on IQ, obesity and blood pressure? Evidence from comparing high-income and middle-income cohorts. *Int J Epidemiol.* 2011;40(3):670–680.

8. Davey Smith G. Assessing intrauterine influences on offspring health outcomes: can epidemiological findings yield robust results? *Basic Clin Pharmacol Toxicol*. 2008;102(2):245–256.
9. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology*. 2010;21(3):383–388.
10. Davey Smith G. Negative control exposures in epidemiologic studies [letter]. *Epidemiology*. 2012;23(2):350–351.
11. Davey Smith G. Use of genetic markers and gene-diet interactions for interrogating population-level causal influences of diet on health. *Genes Nutr*. 2011;6(1):27–43.
12. Heijmans BT, Mill J. The seven plagues of epigenetic epidemiology. *Int J Epidemiol*. 2012;41(1):74–78.
13. Relton CL, Davey Smith G. Is epidemiology ready for epigenetics? *Int J Epidemiol*. 2012;41(1):5–9.
14. Lewington S, Whitlock G, Clarke R, et al. Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. Prospective Studies Collaboration. *Lancet*. 2007;370(9620):1829–1839.
15. Kaptoge S, Di Angelantonio E, Lowe G, et al. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. Emerging Risk Factors Collaboration. *Lancet*. 2010;375(9709):132–140.
16. Wensley F, Gao P, Burgess S, et al. Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data. C Reactive Protein Coronary Heart Disease Genetics Collaboration (CCGC). *BMJ*. 2011;342:d548.

George Davey Smith¹, Matthias Egger², and Shah Ebrahim³
(e-mail: kz.davey-smith@bristol.ac.uk)

¹ *MRC Centre for Causal Analyses in Translational Epidemiology, School of Social and Community Medicine, University of Bristol, Bristol BS8 2BN, United Kingdom*

² *Institute of Social and Preventive Medicine, University of Bern, CH-3012 Bern, Switzerland*

³ *South Asia Network for Chronic Disease, Public Health Foundation of India, New Delhi 110 016, India*

DOI: 10.1093/aje/kwt104; Advance Access publication: June 21, 2013