



Editorial

The future of epidemiology: methods or matter?

Shah Ebrahim¹, Jane E Ferrie² and George Davey Smith^{3*}

¹London School of Hygiene and Tropical Medicine, London WC1E 7HT, ²School of Social and Community Medicine, University of Bristol, Bristol BS8 2BN, ³MRC Integrative Epidemiology Unit, University of Bristol, Bristol BS8 2BN

*Corresponding author. E-mail: ije-editorial@bristol.ac.uk

In our first editorial after taking over responsibility for the content of the *IJE* in 2001, we illustrated some of the successes and failures of epidemiology, and asked the question whether epidemiology would progress or decay in the 21st century?¹ Embracing new opportunities for aetiological understanding provided by the human genome project was one strand of our strategy for making the *IJE* relevant – but viewed through the Geoffrey Rose lens that determinants of disease rates in a population and influences on individual susceptibility need not be the same.^{2,3} We were concerned that the ‘big picture’ – health and disease are as much social as biological phenomena – should remain in view. Indeed, the role for genetic discovery seemed to be much more attuned to the discovery of mechanisms of disease and resulting pharmacological advances. We questioned methodological developments in modern epidemiology which seemed to lack appetite for engaging with the wider uses of epidemiology proposed by Jerry Morris in the 1940s and 50s.^{4–13} A formal evaluation of the impact, positive and negative, of the increased methodological refinement of epidemiology, requested in our 2001 editorial, has yet to be submitted to *IJE*. However, in this the final issue of the journal under our editorship, there are major contributions on causal thinking in epidemiology which illuminate what modern epidemiology can and cannot do.

In 2001, what did we want to do? Up to that point, papers published in the *IJE* had very largely consisted of original manuscripts reporting new data. We noted that the most important paper that the journal had published in its lifetime was one of the few that were not of this type: an invited talk at an International Epidemiology Association meeting (with few references), beautifully and economically summarizing

Geoffrey Rose’s big idea alluded to above, with the catchy title ‘Sick individuals and sick populations’.² Widespread understanding and application of previous epidemiological research was generally lacking, so in our first issue we published a reprint of a 1943 German case-control study on smoking and lung cancer.¹⁴ With its accompanying commentaries,^{15–19} this was the first of the ‘Reprints and Reflections’ which were to become an important *IJE* staple over the next 16 years. Later in 2001 it was the turn of the Rose paper from only 16 years previously (and with rather pleasing symmetry we are now the same distance in time from its reprint).²⁰ Reflecting on the prescience and influence of early contributions to the analysis of genetic influences on human phenotypes by Richard Lewontin^{21,22} or to metabolomic phenotyping by John Moreton,²³ as well as showing how bright minds came to wrong conclusions about the causes of cholera, peptic ulcer and AIDS, can inform future research and act as an immunization against repeating errors. But our innovations were not just about history. From the start we expanded to include editorials, commentaries, point-counterpoint debates, themed issues, theory and methods, and diversions. Over the next 15 years, many other types of article were added: Cohort Profiles, **Health and Demographic Surveillance Systems (HDSS) Profiles**, Data Resource Profiles, Software Application Profiles, Education Corner, Cochrane Corner, Photoessays, Book Reviews, Special Issues and Supplements, together with blogs, press releases and twitter; all brought into being through the enthusiasm, critical eyes and hard work of what grew from 15 associate editors in 2001 to become a group of 45 editors and three editorial staff in 2016.

The question we posed in 2001¹ was whether a bright young scientist who wanted to make a contribution to

population health should build a career in epidemiology or head for the nearest polymerase chain reaction thermal cycler? It is gratifying to see that, in many parts of the world, epidemiology departments have their own genetic, epigenetic and -omics facilities so the question, happily, is now moot. A question that remains of importance is whether the focus on genetic and -omics epidemiology and new methods of causal inference are freezing out applications of epidemiology in health services, public health and clinical medicine.⁷

In 2001 we discussed the prospects for epidemiology in the 21st century under two headings, and then reflected on how the *IJE* could attempt to remain of relevance. We now revisit the same two headings, and conclude with reflections on the current environment for publishing in epidemiology, in many ways even more challenging than existed at the turn of the century.

Genes—explaining all or none of it?

Early successes in infectious disease epidemiology were extended in the 20th century with links established between tobacco and lung cancer, air pollution and winter deaths, and the identification of many cardiovascular disease risk factors, in which epidemiologists were not only involved in identifying causes but also in evaluating methods of prevention. Concerned that ‘big data’, phenomenology and personalized medicine were luring epidemiology away from its traditional purposes, Lew Kuller, writing in 2015,²⁴ noted that ‘Epidemiology is now defined as the collection of large sample sizes and measurement of numerous variables from stored samples to facilitate estimation of disease risk over time’. Although personalized medicine envisaged by Francis Collins, director of US National Institutes of Health, as genetically based, individualized preventive medicine,²⁵ has not gained general traction for diagnosis or screening due to the small effects of most common genetic variants on the risk of common diseases,²⁶ genetic epidemiology has become a core function of most academic departments. It has also laid the path for precision medicine, defined by the National Institutes of Health as ‘an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person’.²⁷ Precision (or stratified) medicine, the form now promoted by Collins,²⁸ provides another opportunity to capitalize on the massive investments in the human genome project, low-cost whole-genome sequencing and more recent -omics research. Precision medicine builds on proof-of-concept studies that have identified specific genotypes among cancer patients and indicate biological pathways and treatments. For example, in chronic

myeloid leukaemia presence of the Philadelphia chromosome somatic anomaly, the *BCR-ABL* fusion gene, produces an abnormal tyrosine kinase protein, and a kinase-inhibitor was found that improved survival dramatically.²⁹ This relatively old example remains a ‘pin-up’ case, but slowly others have emerged. For common diseases, however, the situation is less clear. For example, results from observational studies which supported using genotypes to guide anticoagulation treatment,^{30,31} were not confirmed in three large randomized trials that found genotype-informed algorithms no better than a clinical algorithm in guiding initiation of warfarin treatment.³² Balancing the power of large-scale studies to produce reliable data applicable to groups against the common-sense notion that individuals should be treated according to their very specific characteristics remains, after nearly two centuries, the central conundrum of ‘evidence-based medicine’,³³ and was the focus of another of our historical reprints.^{34–38}

A decade of Mendelian randomization

Whereas the clinical enterprise of precision (or stratified) medicine may gain traction over the next decade, using genetic variants to explore environmentally modifiable (rather than genetic) causal pathways – Mendelian randomization – has, slowly, emerged as a successful approach to harnessing new knowledge gained from the human genome project. Since our original expositions on Mendelian randomization,^{39,40} numerous methodological advances have been made. These range from multiple demonstrations of the plausibility of the basic concept that genetic variants (which can proxy for a potentially modifiable exposure) are essentially unrelated to confounding factors,⁴¹ to the extension of Mendelian randomization into the hypothesis-free resolution of causal directions in correlated networks.^{42,43} This extensive methodological development has been undertaken in response to the challenges of new substantive applied questions and increasingly detailed genetic data and has enabled (and continues to enable) more sophisticated questions to be answered using the framework of Mendelian randomization.

Mendelian randomization has provided definitive answers to both public health and clinical practice questions. For decades public health has struggled with defining whether moderate alcohol consumption is beneficial or harmful for cardiovascular disease. In observational studies, people drinking light-to-moderate amounts of alcohol experienced lower rates of cardiovascular disease⁴⁴ and had lower levels of many cardiovascular risk factors than non-drinkers or heavy drinkers^{45–49} — sometimes termed a ‘J’- or ‘U’-shaped association. There was also evidence that all-cause mortality is lower among men drinking 34 units/week or less and women drinking 16 units/week or less.⁵⁰

These non-linear associations have confused alcohol public health policy, as a ‘safe’ level of drinking for cardiovascular health is consistent with moderate consumption, a level of consumption potentially hazardous for other health outcomes. In a review first published in the *IJE* in 1984 and reprinted in 2001, which reported that moderate drinkers from diverse populations have a lower coronary heart disease (CHD) risk than abstainers, Michael Marmot postulated that abstainers were likely to differ from moderate drinkers in a number of ways.⁵¹ In addition to confounding, such associations are also potentially unreliable because of reverse causation (sick people are told to stop drinking) — situations in which Mendelian randomization can provide alternative means of examining associations. The Alcohol-*ADH1B* consortium has conducted Mendelian randomization analyses showing that people carrying the rs1229984 A-allele of the *ADH1B* gene (a genetic variant associated with lower levels of regular alcohol consumption and a lower risk of alcohol dependence) have a more favourable cardiovascular profile and a reduced risk of coronary heart disease. The researchers concluded ‘Reduction of alcohol consumption, even for light to moderate drinkers, is beneficial for cardiovascular health’.⁵² A development of standard Mendelian randomization analysis, testing for non-linear causal effects, found no strong evidence of U-shaped associations between alcohol consumption and cardiovascular risk factors.⁵³ This and other new evidence appears to have influenced alcohol guidance in the UK to promote lower levels of consumption.⁵⁴ In a similar fashion, high-profile epidemiological publications suggested that moderate alcohol consumption among pregnant women was associated with optimal offspring outcomes, superior to those of non-drinking or more heavily drinking mothers.⁵⁵ Comparison of behavioural and socioeconomic measures by drinking category

for women in the study confirmed Marmot’s hypothesis that abstainers were indeed very different to those who were light drinkers in pregnancy,(Figure 1) and a paper in the *IJE* which applied Mendelian randomization to the question reached the conclusion that any alcohol consumption was detrimental,⁵⁶ a position now recapitulated in UK guidelines (Box 1).

In clinical medicine, Mendelian randomization has contributed to evaluating pharmacological strategies for preventing cardiovascular disease. For example, raised circulating C-reactive protein (CRP) is strongly associated with cardiovascular disease (CVD) and with recurrence of CVD events, suggesting that searching for drugs that lower CRP could provide new means of prevention.⁵⁷ Mendelian randomization studies have now shown convincingly that CRP is not causally related to cardiovascular disease,⁵⁸ thereby avoiding the cost and time of a likely futile evaluation of anti-CRP drugs.

The first cholesteryl ester transfer protein (CETP) inhibitor, torcetrapib, raised high-density lipoprotein (HDL)-cholesterol but also, unexpectedly, increased blood pressure and cardiovascular events.⁵⁹ A Mendelian randomization study was conducted in which associations between variants in the CETP gene, lipid levels and blood pressure were compared with the pharmacological effects. The CETP gene variant that mimicked the torcetrapib effect on lipids, but did not raise blood pressure, indicated that the drug caused an off-target effect, so searching for other CETP inhibitors without such off-target effects would be worthwhile.⁶⁰ However, Mendelian randomization studies indicate that genetic variants that raise HDL cholesterol are not associated with lower risk of cardiovascular disease,⁶¹ predictions borne out by clinical trials of CETP inhibitors mounted by Eli Lilly, Roche and Pfizer.⁶² Merck has continued to support its trial, expected to report in 2017,⁶³ of a CETP-

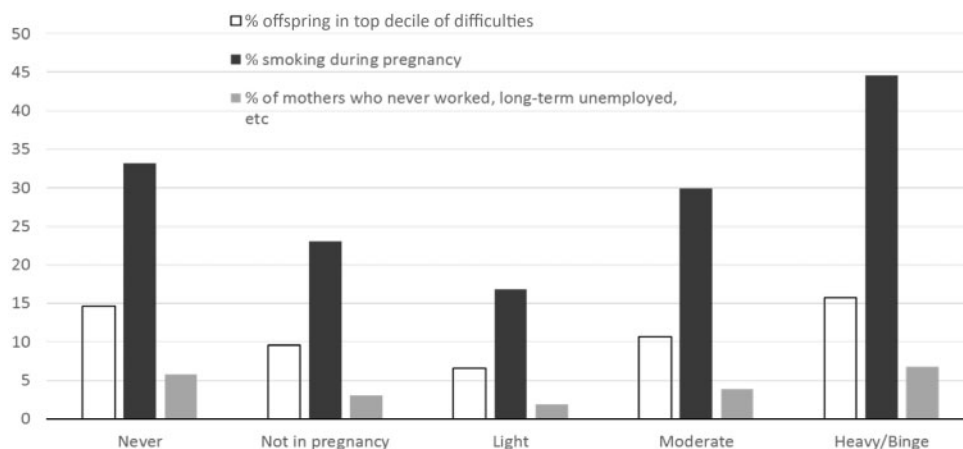


Figure 1. Behavioural difficulties in offspring and the behavioural and socioeconomic characteristics of their mothers, by drinking category. Figure based on data provided in reference 55.

Box 1. Mixed messages on alcohol consumption in pregnancy; which set of 'drinking advice' should pregnant woman follow?

The image displays three news articles related to alcohol consumption during pregnancy. The top article is from 'THE TIMES OF INDIA' with the headline 'Mom's light drinking doesn't harm baby'. The middle article is from 'BBC NEWS HEALTH' with the headline 'Light drinking no risk to baby, say researchers'. The bottom article is from 'ABC News' with the headline 'Pregnant women told to ignore drinking advice'. Each article includes a small image of a pregnant woman and text discussing the findings of a study and official advice.

inhibiting agent that produces a more substantial low-density lipoprotein (LDL)-cholesterol lowering effect than other such agents. However no company would, in the post-Mendelian randomization age, develop a drug solely based on its HDL-cholesterol raising effects.

Statins increase the risk of type 2 diabetes, which may be a specific effect of inhibition of HMG-coenzyme A reductase (*HMGCR*).⁶⁴ Trials of a new class of cholesterol-lowering drug, proprotein convertase subtilisin-kexin type 9 (*PCSK9*) inhibitors,⁶⁵ hoped not to increase diabetes, are underway. However, Mendelian randomization suggests that the relative effects on cardiovascular events and diabetes will be similar for both *PCSK9* and *HMGCR* inhibitor.⁶⁶ The pharmaceutical industry will increasingly undertake Mendelian randomization studies to improve discovery of suitable targets for development, which will hopefully reduce the huge financial losses due to failed phase 3 trials.⁶⁷

Saved from irrelevance by new methodologies?

Representativeness and collider bias

The National Institutes of Health's Precision Medicine Initiative Cohort, slated to enrol a million adults, will recruit a highly non-representative subset of the population – essentially volunteers. As it is often claimed that examining genotype and disease phenotype associations should be robust to any selection biases, does this matter? A recent study, which found that women and children at higher

genetic risk for schizophrenia and related phenotypes were less likely to participate in a large longitudinal study, suggested that such non-participation might introduce such biases.⁶⁸ In a point-counterpoint in this journal, Ken Rothman and colleagues argued that representativeness should be avoided on the grounds that scientific inference benefits from having tightly defined, highly compliant participants.⁶⁹ Although the authors concede that representativeness is required for the practical goal of applying knowledge to populations, they also claim: 'Surveys of opinions, of the prevalence of disease, of habits or of environmental exposures may be informative, but they are not science in the same way that causal studies about how nature operates are science'.⁶⁹ The elevation of causal studies as the science of epidemiology is debated extensively in this issue of the *IJE*. Commentaries on Rothman and colleagues' paper were broadly supportive, although an extensive independent review of this issue was critical of their position.⁷⁰ Richiardi and colleagues in their commentary drew a distinction between intentional (e.g. studying smoking in British doctors to reduce potential confounding by lifestyle factors) and unintentional non-representativeness (e.g. low response rates in recruiting to cohort studies), but considered them to be broadly equivalent in terms of threats to the validity of a study.⁷¹ They argued that the exposure-outcome association might be biased if the exposure of interest and another risk factor for the disease were associated with the probability of inclusion – that is, collider bias⁷² – but presented reassuring evidence from

Monte Carlo simulations and an empirical study that the size of such bias would likely be small.⁷¹

In order to avoid the situation where commentaries on Rothman and colleagues' point-counterpoint article became a one-sided love-in, the editors were forced to give examples of where non-representativeness could be an important cause of bias in exposure-outcome associations.⁷³ In a rebuttal, Rothman and colleagues suggested that we were misguided to assume that a well-designed and conducted randomized clinical trial (RCT) gave the right answer.⁷⁴

It is likely that analyses using large-scale biobank resources comprising volunteer participants (e.g. UK Biobank recruited about 5% of those invited) will generate many spurious findings. Confounding of the exposure of interest with lifestyle or socioeconomic factors, and reverse causation whereby disease status influences exposure rather than vice versa, may invalidate causal interpretations of observed associations. Using data from the British Women's Heart & Health Study we examined 4560 pairwise correlations of non-genetic variables. Of these, at the 1% significance level 46 would be expected to be associated by chance. However, we found associations between 2036 (45%) of them to be 'statistically significant' (not a term of which we approve⁷⁵), giving an observed to expected ratio of 44, P for difference $O:E < 0.000001$.⁴¹ The associations we detected included that hormone replacement therapy (HRT) apparently 'protected' against CHD,⁷⁶ that vitamin C and vitamin E 'protected' against CHD,^{77,78} that antioxidant vitamins apparently improved lung function; all findings from previous observational studies that had led to large-scale RCTs that suggested the associations were non-causal. Although we did not separately publish all these findings we did utilise some of them to illustrate the problems of observational epidemiology.⁷⁶⁻⁷⁸ Indeed many of the findings from other large-scale observational studies, such as the Health Professionals Follow-up Study, Physicians Health Study and Nurses' Health Study cohorts⁷⁹⁻⁸³ have not had their findings confirmed in large RCTs when these have been conducted. Of course it is the most promising leads from observational epidemiology that get taken forward to large-scale RCTs, it is extremely unlikely that the hundreds, if not thousands, of findings from these cohorts which have not been followed up with RCTs are somehow magically reliable, whereas the apparently most reliable ones that were followed-up with RCTs are spurious. There often appears to be no reckoning in epidemiology, however. Consider, for example, findings on vitamin E supplementation – with the exposure measured in the observational study (putting vitamin E supplements in to your mouth and swallowing them) identical to the exposure then tested without success in RCTs. Rather than go back and try to account for why studies have got things wrong, in general the response is to move

forward and churn out more such findings, that then obtain the same media splash. Biobanks even larger in scale than the Harvard studies unlikely have yet greater possibilities for generating spurious findings to fill medical journals and newspapers.

An important distinction in this discussion of representativeness is the potential clear-cut effects of selection into a target population, the response rates in a study based on that target population, and attrition rates among the responders (in long-term cohort studies). Selection into a target population was discussed in a short 1911 article by Arthur Pigou on the relationship between parental alcoholism and adverse offspring development. Pigou's critique targeted high-profile analyses led by Karl Pearson that used data from particular deprived geographical regions. Pigou argued that comparisons between the offspring of alcoholic and non-alcoholic parents would be biased: the former were better-off as they were able to pay rent and buy alcohol, whereas the latter only had to be able to pay rent, but both found themselves in the same miserable target areas.⁸⁴ Simulation studies indicate that even modest influences on selection into or attrition from a study can generate biased and potentially misleading estimates of both phenotypic and genotypic associations.⁸⁵

Traits that are entirely unrelated in the general population may appear to be correlated in selected samples as a result of collider bias, if both traits influence initial selection into the study population, or loss to follow-up in the analysed sample. Genetic variants are, in general, unrelated to confounders. When many variants are combined into polygenic scores that are associated with a phenotype, an association between the phenotype and participation will allow the score to be more strongly related to participation than is each individual variant.⁶⁸ This, in turn, can potentially lead to bias. Studies using polygenic scores, genome-wide allelic scores⁸⁶ and whole-genome genetic correlations (including linkage disequilibrium regression)^{87,88} in highly selected samples are most at risk of producing biased and potentially misleading results.

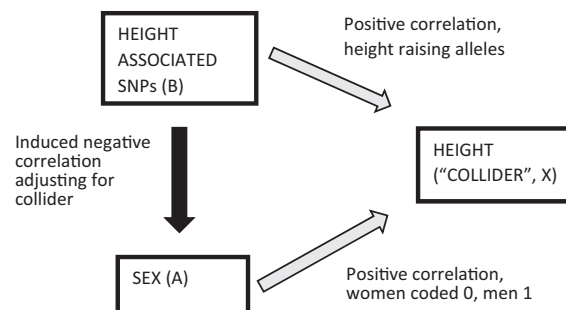


Figure 2. A schematic representation of a biologically spurious association between SNPs associated with height (B) and sex (A) can be achieved by adjusting for height which acts as a collider (X). Adapted from reference 89 with permission from Elsevier.

A recent toy example using UK Biobank data showed how a biologically spurious association between single nucleotide polymorphisms (SNPs) associated with height (B) and sex (A) could be achieved by adjusting for height which acts as a collider (see Figure 2).⁸⁹ Less contrived examples of this and related problems are beginning to appear.^{90,91} Thus the aforementioned study with genetic data at birth demonstrated that common genetic variants associated with schizophrenia were also associated with non-participation, non-completion of questionnaires and non-attendance at data collection throughout childhood and adolescence.⁶⁸ Analyses of schizophrenia-related phenotypes as outcomes may be biased by the non-random missingness of these phenotypes in the study population – for example, an implausible inverse association between the polygenic score for schizophrenia and reported psychosis-like symptoms is seen among the respondents.⁹²

Causal inference and DAGs

Despite eloquent presentations of the purpose and uses of epidemiology over the past century,^{5,93,94} concerns about its direction have been voiced repeatedly.^{7,95–98} Earlier discourses highlighted the role of epidemiology in improving public health and providing a scientific basis for clinical practice. Much debate now centres on the overemphasis on identifying causes of disease as the sole purpose of ‘modern

epidemiology’, and with an obsession with ever more complex statistical methodology. In this issue we have brought together articles written by some of the more influential epidemiologists involved in advancing causal methodology, and a series of critical commentaries to develop and further the debate about the uses of epidemiology.^{99–111}

Arguments for and against modern ideas about causal inference revolve around the ways in which causes should be defined. The potential outcomes approach (related to counterfactual thinking) can be seen as too rigid and too far removed from many of the complex ‘dirty’ problems (e.g. social inequalities, racism, ecological changes) of social epidemiology. If a potential ‘cause’ cannot be manipulated, is it sensible to disregard it, relegating it to the ‘not suitable for epidemiology’ category? The use of directed acyclic graphs (DAGs) may, if properly constructed, aid causal thinking and help plan relevant analyses.¹¹² But DAGs and analyses may increasingly be constructed by computer programs without sufficient application of biological/pathological knowledge, and it is not clear whether the resulting DAGs aid analysis or drive them, or whether they constrain what can be analysed and understood (see Figure 3).

The authors responsible for Figure 3 conclude ‘BMI and physical activity in early childhood are associated with atopic sensitization, atopic dermatitis and asthma in later childhood’.¹¹³ However, the DAG does not provide a comprehensive picture. For example, it does not include

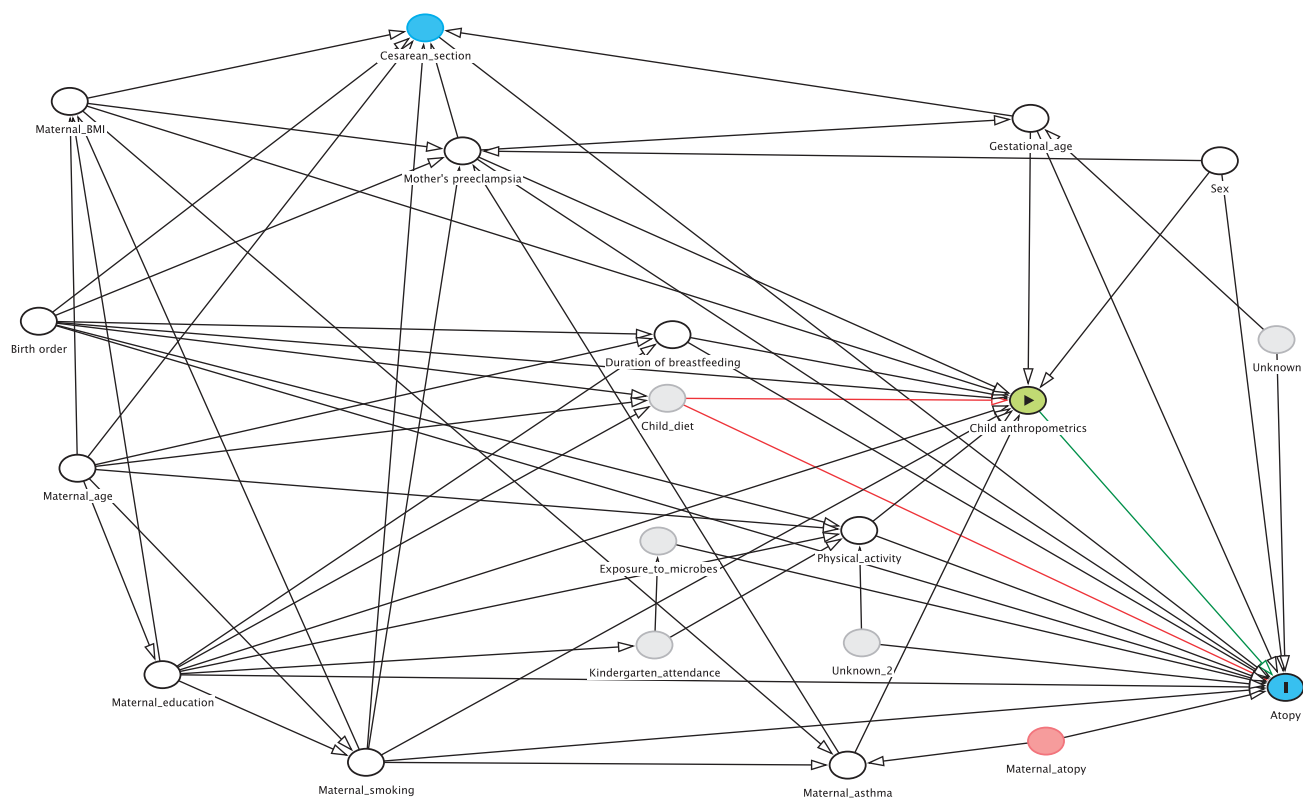


Figure 3. A DAGitty diagram to explore associations of body mass index (BMI) and physical activity with atopy.¹¹³ Licensed under CC BY 4.0.

paternal factors,¹¹⁴ ethnicity,¹¹⁵ respiratory infections¹¹⁶ or socioeconomic position,¹¹⁷ nor distinguish different atopy phenotypes.¹¹⁸ DAGs, however elegant and expansive, are unlikely to improve causal inference without prior knowledge and, ultimately, need the causal anchors – proposed by Sewall Wright in his work on path analysis in the 1920s – to make reliable causal interpretations.¹¹⁹ The massive and still growing knowledge available on the causal effects of germline genetic variants on phenotypes provide such anchors, and in some circumstances enable better exploitation of observational data using Mendelian randomization approaches to causal inference. Investigators using the machine algorithms now built into DAG programs should heed Siddhartha Mukherjee's warning: 'The medical revolution will not be algorithmized'.¹²⁰

Does a more pluralistic approach – e.g. inference to the best explanation¹²¹ – to identifying causes help more? How will new proposals to develop a more formal approach to triangulation of findings, defined as 'the practice of strengthening causal inference by integrating results from several different approaches, where each approach has different (and assumed to be largely unrelated) key sources of potential bias',¹⁰⁴ contribute to how we carry out epidemiology in the future?

Population health sciences: epidemiology in smarter clothes?

Population health sciences is a growing field, popular with graduate students and doctoral aspirants, defined as 'the study of the conditions that shape distributions of health within and across populations, and of the mechanisms through which these conditions manifest in the health of individuals'.¹²² This quotation is taken from a primer by Katherine Keyes and Sandro Galea, for students of this new discipline. Keyes and Galea's primer is dedicated to Geoffrey Rose's aforementioned seminal work on prevention,⁹⁶ in which he articulated the importance of two premises: (i) distinguishing the causes of susceptibility from the causes of incidence – sick individuals and sick populations; and (ii) that shifting the population distribution of adverse exposures downwards will prevent more disease than focusing on those at the top end (high risk) of the distribution.² Their definition of population health science clearly overlaps with most definitions of epidemiology, e.g. 'the study of the distribution and determinants of health-related states or events (including disease), and the application of this study to the control of diseases and other health problems'.¹²³

Is there a need for this new discipline? The rarefied environment of modern epidemiology has, in some schools of public health, focused so much on causal inference using counterfactual reasoning and on ever more complex

statistical methods, that there is little room for working across disciplines – anthropology, biology, demography, economics, genetics, medicine, politics, public health, psychology, sociology, etc. – and working on problems of importance – ageing, climate change, conflict, development, emergent infections and pandemics, equity, health and social care, global health, etc. In their short primer, Keyes and Galea promote exciting ideas and practical ways of integrating methods drawn from different disciplines to tackle problems of consequence. This issue of the *IJE*, bringing together articles on causal inference, may help in identifying the common ground between epidemiologists who appear to have taken polar positions and, we hope, like Keyes and Galea's primer, will be useful for building theory and practice in epidemiology.

The need for research cutting across disciplines is exemplified by a systematic review of studies evaluating the long-term economic impacts of the deworming children in poor countries (see 'Re-analysis, re-appraisal and re-interpretation of data' below). Reviewing the research from epidemiological and public health perspectives, rather than solely from an economic viewpoint, results in rather different conclusions. Supporting population health science initiatives in departments of epidemiology may be part of the solution to achieving relevance and traction on our current (and growing) health problems.

Big data

Big data can be defined in terms of size, complexity, manipulability and management,¹²⁴ and in the context of health sciences ranges from electronic patient records¹²⁵ to open-access genetic and phenotypic datasets. In the latter category, UK Biobank is the current front-runner and has become an extremely valuable resource. However, ease of access requires a careful approach to data analysis and interpretation if spurious findings due to selection bias, confounding or just straight-forward data torture are to be avoided. An important step in improving causal inferences will be for investigators to publish their analytical code along with their paper, to enable replication and additional analyses to be conducted. Other sources of big data also have their advantages and drawbacks. Big data derived from hospital episode statistics and from routine contacts in primary care are widely used to explore health care delivery, effectiveness and variation in outcomes. For example, the Clinical Practice Research Datalink (CPRD) provides an ongoing primary care database of anonymized medical records from general practitioners, with coverage of over 11.3 million patients from 674 practices in the UK.¹²⁶ This database was used to examine the association between measles, mumps and rubella (MMR) vaccine and

autism, and successfully debunked the idea that MMR caused autism.¹²⁷ In England a large-scale health policy change – ‘the 7-day NHS’ – was initiated on the basis of an analysis of hospital episode statistics (14.8 million hospital admissions) that purported to show excess mortality at weekends.¹²⁸ The investigators refused to share their analytic code on grounds of research governance,¹²⁹ thus prohibiting other investigators from exploring the robustness of their findings using sensitivity analyses and other approaches. It is only a matter of time before all research papers will share their analytic code for the common good.¹³⁰ The extra effort involved in providing shareable analytic code should be seen as time well spent and would avoid the issues that arise when analytic code is lost for future interrogation.^{131,132}

Re-analysis, re-appraisal and re-interpretation of data

In low- and middle-income countries (LMICs), issues of analysis and interpretation of data can have far reaching consequences. Deworming – mass medication of children and women of child-bearing age in LMICs – is a massive initiative involving the World Health Organization and the ‘Deworm the World’ campaign.¹³³ The rationale is simple – a 6-monthly pill will improve nutritional status, cognition and school attendance and thereby promote economic development. The underpinning evidence includes an analysis by Edward Miguel and Michael Kremer of a study conducted in Kenya in 1998–99,¹³⁴ augmented by data from a study from 1910,¹³⁵ and an unpublished report of a follow-up of the original Kenya study.¹³⁶ The original data and analytic code were made available by Miguel and Kremer, and an independent team of investigators undertook a replication and re-analysis, which was published in the *IJE* in 2015.^{137,138} The re-analyses found a small effect of deworming on school attendance but no clear effect on examination performance.¹³⁹ Perhaps the major triumph here is not so much the findings themselves but that the data and code were made available for independent scrutiny – an approach that is both inexpensive and quick compared with mounting new trials.¹⁴⁰

An updated Cochrane review of randomized controlled trials of deworming found no strong evidence of benefits for nutritional status, haemoglobin levels, school attendance or examination performance,¹⁴¹ and a mega-trial of two million children in India failed to demonstrate benefits from deworming on mortality or weight gain.¹⁴² A further piece of evidence is a review by Sophie Jullien and colleagues of the unpublished economic studies evaluating the long-term health, schooling and economic development effects of the original trials in Kenya and Uganda,¹⁴³

published in this issue of the *IJE*. Disseminated online as working papers without formal peer review and publication, two of the three papers have gone through multiple iterations. Nonetheless they are widely cited as claiming benefits for health and economic development, by advocates of deworming. However, according to Jullien and colleagues, they do not appear to stand up to an epidemiological and systematic review approach, and the authors conclude: ‘In the context of reliable epidemiological methods, all three studies are at risk of substantial methodological bias. They therefore help in generating hypotheses, but should not be considered reliable evidence of effects.’¹⁴³

Are downstream benefits, such as economic development, feasible in light of no clear benefits for upstream outcomes?¹⁴⁴ Despite the growing evidence showing that deworming has little to offer at a population level, it remains popular with the World Health Organization¹⁴⁵ and the ‘Effective Altruism’ movement, who have claimed that deworming is more effective than providing textbooks or teachers in Kenya.¹⁴⁶ A so-called ‘worm wars’ for and against continuing current deworming policy in light of this new evidence have been declared, with considerable media coverage.¹⁴⁷ The commentators comprise the authors of the original working papers,^{134–136, 148–150} GiveWell¹⁵¹ (an organization that assesses evidence on interventions supported by charities) and other scientists with methodological interests.^{152–155} Not surprisingly, there are points of difference and of emphasis about the analyses by Jullien and colleagues, many of which are covered in the eight commentaries on their review.^{148–155} However, there is some agreement that the evidence base for deworming is sparse, and further evaluations of its long-term developmental and economic effects would be helpful.

The next 16 years. Where is (publishing) epidemiology going?

Missing the obvious: social inequalities in health

The 1980s were the decade when social inequalities became a major preoccupation with epidemiologists. Commissioned in the UK under a Labour administration, the Black Report¹⁵⁶ on social inequalities in health was published by the newly elected conservative Thatcher government on a national holiday in August 1980, in the hope it would be ignored. Instead it became an inspiration for a generation of epidemiologists and social scientists around the world, and social inequalities were described for almost every measure of health and disease across countries rich and poor. However, tackling social inequalities in health

from the epidemiological viewpoint proved difficult. Interventions, such as CVD prevention programmes that are focused on individual behaviour change, were differentially taken up by more advantaged groups in the population, thereby exacerbating social inequalities.^{157,158} More recently the social inequalities theme has rather run out of

steam, despite valiant efforts by Michael Marmot and colleagues to keep it high up research and policy agendas worldwide.¹⁵⁹⁻¹⁶¹ This wane in interest has been attributed by some to the increasing interest in genetics and by others to the obvious lack of impact on the underlying drivers of inequality (Figure 4). Indeed, describing trends in terms of absolute and relative changes in risk by socioeconomic position have further confused the question of whether outcomes are getting worse or better for the disadvantaged in any country. Taking Britain as a case study, the absolute rates of decline in CHD in women under age 75 show a marked decline in all social class groups, and a decrease in the difference in absolute mortality rates between them (Figure 5).¹⁶²

By contrast, when the relative rates of decline are examined (Figure 6)¹⁶² it is clear that the more disadvantaged sections of the population have actually been doing worse over the past two decades. Progress or not? It seems the social inequalities stream of research and activism (in which many epidemiologists, to their credit, played a role) could not withstand the much stronger political and commercial forces that have dominated public policy in most countries over the past two decades. With a newly appointed regressive government in the USA, and with India, Russia and much of Europe and Latin America following policies that will not support disadvantaged groups, a resurgence of the social inequality agenda may emerge following the likely patterns of health inequalities the one will ensue. Such a movement might want to address such informative questions such as the one posed by Oxford geographer Danny Dorling: ‘Can we afford the rich?’¹⁶³ in order to develop a better understanding of how those in power maintain their grip on society.¹⁶⁴ There is a role for epidemiology here

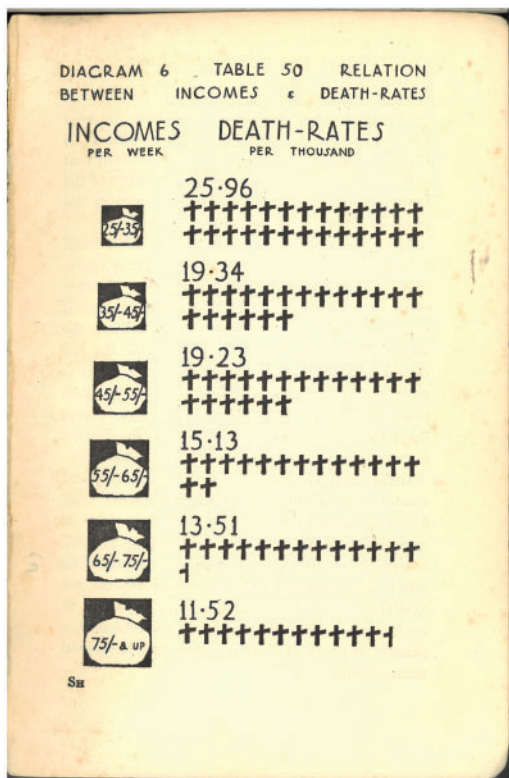


Figure 4. Relation between incomes and death rates. In: Poverty and Public Health by G.C.M. M'Gonigle and J. Kirby, published by Victor Gollancz, London, 1936.

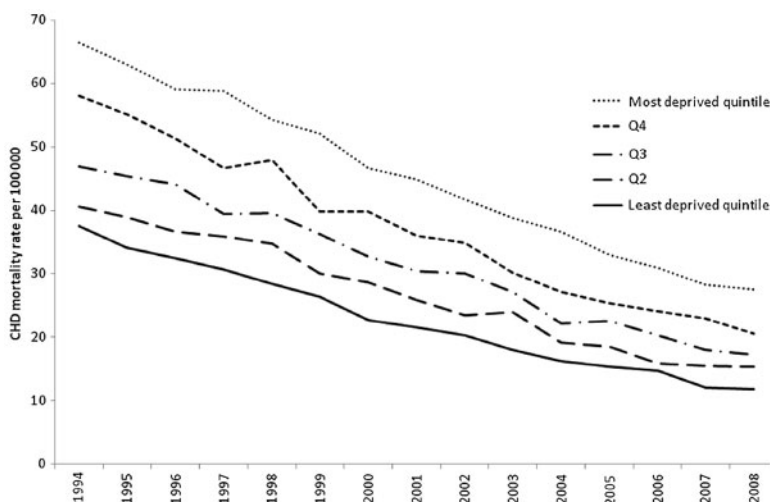


Figure 5. Age-standardized coronary heart disease mortality rates per 100 000, 1994-2008 by quintile of deprivation. Great Britain, women younger than 75 years. Reproduced with permission from BMJ Publishing Group Ltd.¹⁶²

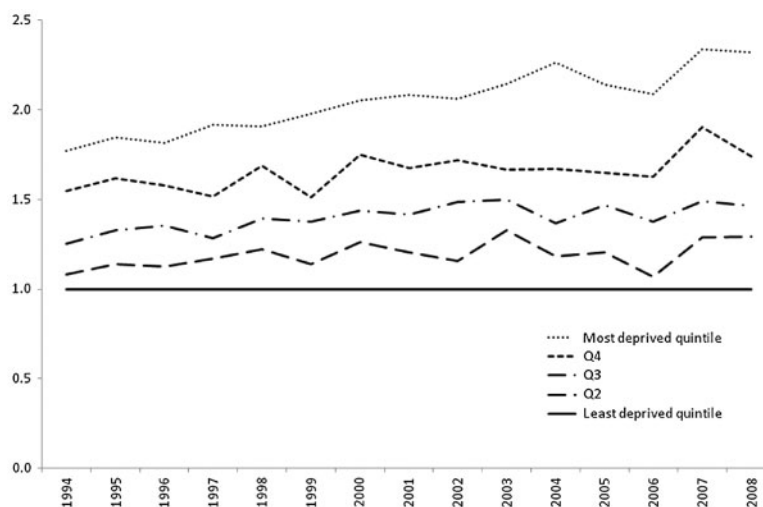


Figure 6. Coronary heart disease mortality rate ratios (least deprived quintile as baseline), 1994 to 2008 by quintile of deprivation. Great Britain, women younger than 75 years. Reproduced with permission from BMJ Publishing Group Ltd.¹⁶²

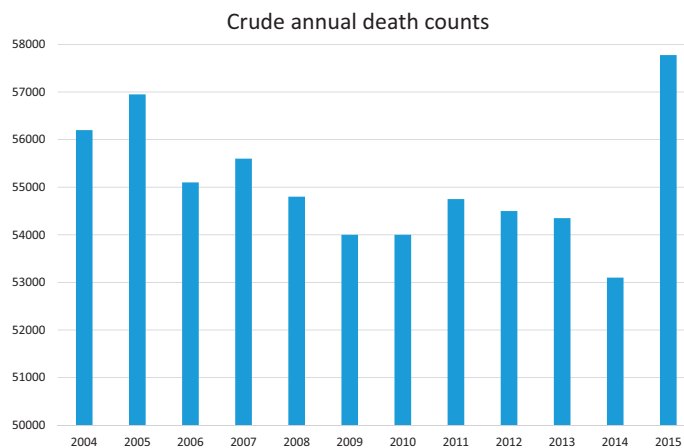


Figure 7. Numbers of deaths per year, 2004-15 Scotland. Reproduced with permission from Danny Dorling.¹⁶⁶

and it goes beyond description to identifying with greater precision which interventions – political, social and health care — have an impact on health and development, and in which contexts. Counterfactual methods focusing on identifying intervention effects might help drive a new wave of social epidemiology focused on reducing social inequalities in health,¹⁶⁵ but care must be taken to ensure that proposals as to what can be done are not constrained by what are considered to be acceptable questions.

Monitoring and surveillance: back to basics

Monitoring and surveillance using epidemiological data are fundamental for assessing the health of populations, but compared with the tantalising prospect of ‘discovery’ research offered by the unravelling of the human genome, epigenetics and -omics technologies, they have become areas of lesser interest. Dorling lamented this phenomenon

at a recent meeting [<http://www.ijeconference.com/live-stream/>] when he asked why epidemiologists were not interested in the 9% increase in deaths that occurred in Scotland in 2015 (see Figure 7), and the stagnation of life expectancy there for the first time in 160 years.

The Global Burden of Disease (GBD) project started over 25 years ago. In the absence of robust mortality statistics from most LMICs, it attempts to provide much-needed information about patterns of disease and risk factors by geographical region and over time. The complex methods used to derive GBD estimates make them opaque and difficult to verify. For example, estimates of Indian trends in blood cholesterol and blood glucose between 1980 and 2008 were not based on nationally representative data but relied on regional studies conducted between 1994 and 2006 (cholesterol) and 1989 and 2007 (glucose).¹⁶⁷ Statistical models filled in the gaps by ‘borrowing’ from other countries in the South Asia region, modelling

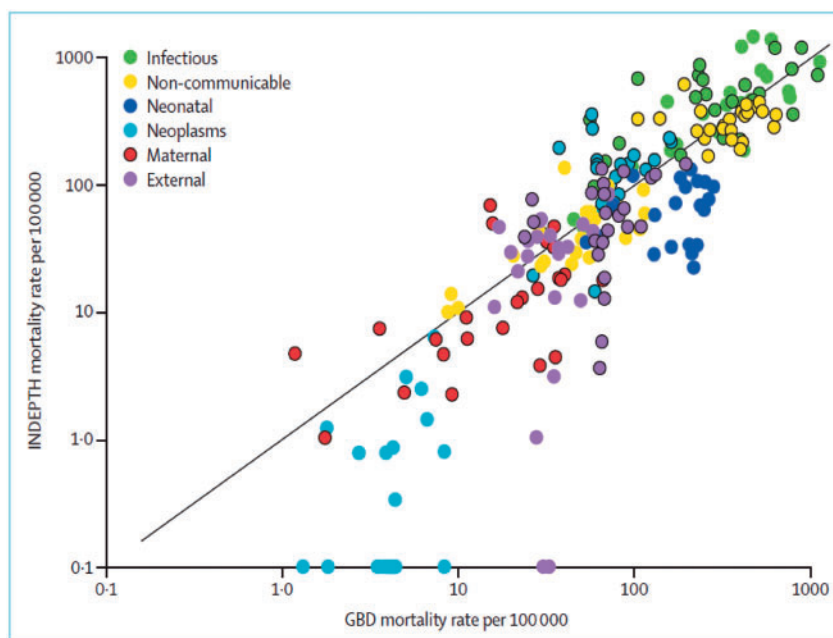


Figure 8. Concordance correlation between GBD and INDEPTH cause-specific mortality findings in 13 low-income and middle-income countries by six major cause of death categories. Each point represents one country, cause category, age group and 5-year period. The diagonal black line represents equivalence. Circles with solid outlines = > 15 years of age. Circles with no outline = < 15 years of age.¹⁷⁰ Reproduced under Creative Commons Attribution (CC BY 4.0).

estimates using income, urbanization and multiple food type availability.^{168,169}

Excluded from the 2013 analyses, previous GBD reports had included cause-specific mortality data from some INDEPTH HDSS sites. This provided an opportunity to make a direct comparison between GBD estimates of cause specific mortality and 'real-world' estimates from verbal autopsy data from the INDEPTH field sites. Overall concordance between the two data sources over 50 causes of death, two age groups and three periods was 0.585, increasing to 0.770 when six major cause categories were used (see Figure 8).¹⁷⁰ Many countries in Africa and Asia are not able to ascertain cause-specific mortality routinely using surveillance methods, so GBD estimates provide the only information to guide policy. Continued investment in and creation of more INDEPTH HDSS sites are essential for providing real country-specific data, to generate data on risk factors, and to provide community 'laboratories' in which interventions and health system programmes can be evaluated and epidemiological training provided.¹⁷¹

Much was written in 2015 about the notable successes and failures of the Millennium Development Goals (MDGs). In an early editorial on the subject in the *IJE*, Alan Lopez, lamented the omission of a measure of health inequalities in the MDGs and the lack of focus on non-communicable disease, as missed opportunities.¹⁷² Secretary General of the United Nations (UN), Dr Margaret Chan, described 2015 as a time for global action. 'During this single year we have the unequivocal opportunity and

responsibility to adopt sustainable development, to restructure the global financial system in line with our needs, and to respond finally and urgently to the challenge of human-induced climate change'.¹⁷³ Sadly, in common with tackling social inequalities in health, global action on the Secretary General's aspirations fell woefully short. Nonetheless, in September 2015 the UN General Assembly established the Sustainable Development Goals (SDGs); 17 universal goals, 169 targets and 230 indicators leading up to 2030. GBD collaborators used indicators from the GBD study in 2015 to set the baseline for the 33 health-related SDG indicators and to examine progress on these between 2000 and 2015. Although some overall progress was documented, as might be expected, greatest progress was seen for the MDG-related indicators.¹⁷⁴ Despite this, the authors observe that gains on the health-related MDG indicators will need to be sustained and, in many cases, accelerated if the ambitious SDG targets are to be achieved.

Global health: equitable authorship

Submissions to the *IJE* from LMIC authors made up only 19% of the total in 2015; a figure that has changed little over the past decade. Each year, we have received papers from high-income country (HIC) authors, using data derived from LMICs that include no authors from the relevant countries. Presumably the investigators believed that no one deserved authorship, even though it is hard to understand how the data were obtained without the

Get me off Your Fucking Mailing List

David Mazières and Eddie Kohler
 New York University
 University of California, Los Angeles
<http://www.mailavenger.org/>

Abstract

Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list.

your fucking mailing list. Get me off your fucking mailing list.

Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list. Get me off your fucking mailing list.

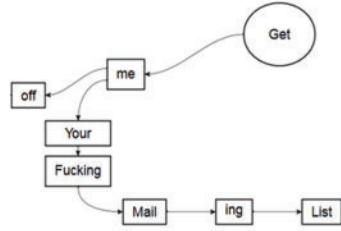


Figure 1: Get me off your fucking mailing list.



Figure 9. Left panel—an article that was actually submitted to an online journal, *International Journal of Advanced Computer Technology*, and was reviewed as ‘excellent’. Reproduced with the kind permission of Eddie Kohler and David Mazières.¹⁸⁰ Right panel—Banksy’s view on what needs to happen, photo courtesy of George Davey Smith.

involvement of LMIC researchers. A variation we observed was token authorship for investigators from LMICs in the middle of the author list – the ‘thinly filled LMIC sandwich’ approach to authorship: problems identified in previous issues of the journal.^{175,176}

In this issue of the *IJE* we publish one of the first studies to explore lead authorship for researchers in LMICs. Using randomized controlled trials of interventions for major infectious diseases (i.e. HIV/AIDS, malaria and tuberculosis) conducted in LMICs as indicator articles, Margaret Kelaher and colleagues demonstrated a 5-fold increase in the number of articles in the post-MDG (2000-12) period than in the 10 years before. However, proportionally the increase was 3-fold greater for HIC first-authorship compared with LMIC first-authorship. LMIC first authorship increased over time for research funded from LMICs, but declined for research funded by US and non-US HIC sources.¹⁷⁷ These data suggest that the Global Health movement currently benefits HIC institutions and investigators at the expense of LMIC researchers – surely an unacceptable form of neo-colonialism?

Is publishing epidemiology going to continue?

Most epidemiological studies are not published in epidemiology journals, partly because epidemiology journals have lower impact factors than general medical and specialty

journals. Furthermore, epidemiological studies often gain wide public attention, especially when they concern the effects of lifestyles on health which provide tantalizing media tweets, help sell newspapers, promote the journals in which they are published and fulfil funders’ requirement for dissemination. Unfortunately, many of these studies massage tiny relative risks into major hazards to health, or produce public confusion by reporting contradictory findings. Anything with sex in the title is a winner; a 2015 *IJE* paper by Niklas Långström and colleagues, on sexual offending,¹⁷⁸ not only resulted in extensive media coverage but also in over 57 000 downloads of the original paper. Ticking so many boxes, it seems likely that this stream of epidemiology will continue for the next 16 years.

The opportunities for web publication of findings, independent of commercial or profit-for-purpose publishers, have been taken up by several scientific disciplines (e.g. physics, astronomy). In medical and related disciplines, publication in print and/or online journals remains the main route for dissemination and discussion of findings, and it seems unlikely that these journals will simply disappear. Richard Smith, former editor of the *BMJ*, tells a story about his early experiences of being an editor [<http://www.ijeconference.com/livestream/>]. The editors of a medical society journal, published by BMJ Publications, asked:

‘What is the added value of the publisher for the academic community? We get grants, do the research and

Box 2. Extract from 'High Windows' by Philip Larkin

I know this is paradise

Everyone old has dreamed of all their lives—
 Bonds and gestures pushed to one side
 Like an outdated combine harvester,
 And everyone young going down the long slide

To happiness, endlessly.

Philip Larkin, 'High Windows' from *Collected Poems*.
 Copyright © Estate of Philip Larkin. Full poem available
 at <https://www.poetryfoundation.org/poems-and-poets/poems/detail/48417>.¹⁸⁴

write the papers; we do the peer review, decide which papers to publish, obtain commentaries and editorials, and deal with letters, appeals and complaints. All of this is at no cost to the publisher but you take a large share of the profits generated by advertising and library subscriptions.'

It is this last issue that explains why medical journals stay in business and medical societies are better off with publishers – the profits that publishers make are shared with the medical societies and provide a major source of income. In the case of the *International Epidemiology Association* (IEA), the *IJE* provides 80% of its annual income, and without the *IJE* the IEA's ability to function would be seriously compromised.

In our first editorial in 2001¹ we applauded the BioMed Central initiative and have since been impressed by the growth of PLoS journals, which attempt to contextualize articles and have managed to engage with many prominent researchers. However, we did not foresee the requirements by major research funders for articles to be open access on publication, nor the massive growth of online-only medical journals. The latter are now making it very difficult to discern which journals are legitimate and which are predatory – that is, established simply to make money and have no or very little editorial oversight or input, or appear to be compiled by machines (see [Figure 9](#)).^{179,180} In collaboration with Oxford University Press (OUP), which publishes the *IJE*, we had planned to launch an online-only journal – *IJE Open* – to provide a forum for descriptive studies from LMICs and to have a strong focus on population health. Unfortunately, these plans did not withstand changes of management within OUP and IEA.

A solution to sifting the digestible from the garbage has recently been provided by the Wellcome Trust, a major UK biomedical research funder. The 'Wellcome Open Research' initiative¹⁸¹ aims to remove the science funded by the Trust from the grasp of publishers and onto its own online platform, which uses services developed by the

F1000 Research platform.¹⁸² Following a basic check by the in-house editorial team, manuscripts are uploaded to the platform with all processing charges covered by the Wellcome Trust. Open peer review is conducted after publication and posted with the manuscript. There is no rejection or acceptance process, authors select their own reviewers and can comment and/or revise their manuscripts in response to the reviews, which are also posted with the paper. Wellcome-funded researchers are not obliged to use Wellcome Open and can still choose to publish papers in traditional journals. Although Wellcome's purpose is to ensure that all findings are in the public domain and not just those that provide strong evidence of association, it is up to Wellcome grantees to decide which results they think are worth sharing using Wellcome Open. A major advance is that underlying data/code will be available to those who wish to reuse it. Other major national and international funders are likely to follow suit, making it quicker for findings to be disseminated, and easier for the end user to find quality research at no direct cost. BioRxiv [<http://biorxiv.org/>] is an earlier initiative providing a pre-print server for life sciences run by Cold Spring Harbor Laboratory. It has grown dramatically from five pages of articles in January 2014 to 67 pages in December 2016. This allows investigators to submit their work, which is then citable, before publication in traditional journals. There is no peer review, editing or formatting of articles, but all are checked for offensive and/or non-scientific content and plagiarism. The Medical Research Council in the UK has recently agreed that articles uploaded to bioRxiv can be cited in grants, reports and curricula vitae.

Time to call it a day

Although we never had the nerve (despite the occasional need) to follow Hunter S. Thompson's editorial approach at *Rolling Stone* magazine: 'What kind of lame, half-mad bullshit are you trying to sneak over on us? Do you take us for a gang of brainless lizards? Rich hoodlums? Dilettante thugs? ... Get your worthless ass out of the piazza and back to the typewriter. Your type are a dime a dozen around here...'¹⁸³ (and this to the distinguished author Anthony Burgess); overwhelmingly, editing the *IJE* has been intellectually stimulating and enormous fun. Of course we regret the IEA's decision to cut our funding to the point where it was no longer viable for us to continue. Nonetheless, we wish the new editorial team under Stephen Leeder all the best in their editorial endeavours to 'un-fuck the system' ([Figure 9](#)). We would also like to take this opportunity to thank our authors and commentators, the members of our editorial board and editorial staff for their unstinting dedication and support – it's been a great ride,

but it is time for us to stop, in the words of Philip Larkin, ‘going down the long slide to happiness, endlessly’ (Box 2).¹⁸⁴ We are no longer young, and don’t deserve such.

Acknowledgments

We are grateful to Nish Chattervedi, Debbie Lawlor, John Lynch and Caroline Relton for helpful comments on this article.

References

- Davey Smith G, Ebrahim S. Epidemiology—is it time to call it a day? *Int J Epidemiol* 2001;30:1–11.
- Rose G. Sick individuals and sick populations. *Int J Epidemiol* 1985;14:32–38.
- Davey Smith G. Epidemiology, epigenetics and the ‘Gloomy Prospect’: embracing randomness in population health research and practice. *Int J Epidemiol* 2011;40:537–62.
- Morris JN. *Health*. London: English Universities Press, 1944.
- Morris JN. *Uses of Epidemiology*. Edinburgh: Livingstone 1957.
- Morris JN. Uses of epidemiology. (Reprint) *Int J Epidemiol* 2007;36:1165–72.
- Davey Smith G. The uses of ‘Uses of Epidemiology’. *Int J Epidemiol* 2001;30:1146–55.
- Loughlin K. Epidemiology, social medicine and public health. A celebration of the 90th birthday of Professor J. N. Morris. *Int J Epidemiol* 2001;30:1198–99.
- Berridge V. Jerry Morris. *Int J Epidemiol* 2001;30:1141–45.
- Marmot M. From Black to Acheson: two decades of concern with inequalities in health. A celebration of the 90th birthday of Professor Jerry Morris. *Int J Epidemiol* 2001;30:1165–71.
- Oakley A. Fifty years of JN Morris’s *Uses of Epidemiology*. *Int J Epidemiol* 2007;36:1184–85.
- Porter D. Calculating health and social change: an essay on Jerry Morris and late-modernist epidemiology. *Int J Epidemiol* 2007;36:1180–84.
- Krieger N. Commentary: Ways of asking and ways of living: reflections on the 50th anniversary of Morris’ ever-useful *Uses of Epidemiology*. *Int J Epidemiol* 2007;36:1173–80.
- Schairer E, Schöniger E. Lung cancer and tobacco consumption. *Int J Epidemiol* 2001;30:24–27.
- Schairer HU. Commentary: In memoriam of my father, Prof. Dr. Med. Dietrich Eberhard Schairer. *Int J Epidemiol* 2001;30:28–29.
- Doll R. Commentary: Lung cancer and tobacco consumption. *Int J Epidemiol* 2001;30:30–31.
- Zimmermann S, Egger M, Hossfeld U. Commentary: Pioneering research into smoking and health in Nazi Germany—The ‘Wissenschaftliches Institut zur Erforschung der Tabakgefahren’ in Jena. *Int J Epidemiol* 2001;30:35–37.
- Proctor RN. Commentary: Schairer and Schöniger’s forgotten tobacco epidemiology and the Nazi quest for racial purity. *Int J Epidemiol* 2001;30:31–34.
- Ernst E. Commentary: The Third Reich—German physicians between resistance and participation. *Int J Epidemiol* 2001;30:37–42.
- Rose G. Sick individuals and sick populations. *Int J Epidemiol* 2001;30:427–32.
- Lewontin RC. Senior Lecture, Annual Short Course on Medical and Experimental Mammalian Genetics, Bar Harbor, Maine, 1968. *Int J Epidemiol* 2016;45:654–64.
- Lewontin RC. Reprint. The analysis of variance and the analysis of causes. *Int J Epidemiol* 2006;35:520–25.
- Moreton JR. Chylomicronemia, fat tolerance, and atherosclerosis. *Int J Epidemiol* 2016;45:1372–79.
- Kuller LH. Epidemiology: then and now. *Am J Epidemiol* 2016;183:372–80.
- Collins FS. Medical and Societal Consequences of the Human Genome Project. *N Engl J Med* 1999;341:28–37.
- Davey Smith G, Ebrahim S, Lewis S, Hansell AL, Palmer LJ, Burton PR. Genetic epidemiology and public health: hope, hype, and future prospects. *Lancet* 2005;366:1484–98.
- The Precision Medicine Initiative Cohort Program. Building a Research Foundation for 21st Century Medicine. Working Group Report to the Advisory Committee to the Director, National Institutes of Health. 2015. http://acd.od.nih.gov/reports/PMI_WG_report_2015-09-17-Final.pdf (15 December 2016, date last accessed).
- Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015;372:793–95.
- Druker B, Talpaz M, Resta DJ *et al*. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N Engl J Med* 2001;344:1031–37.
- Epstein RS, Moyer TP, Aubert RE *et al*. Warfarin genotyping reduces hospitalization rates results from the MM-WES (Medco-Mayo Warfarin Effectiveness study). *J Am Coll Cardiol* 2010;55:2804–12.
- Gage BF, Eby C, Johnson JA *et al*. Use of pharmacogenetic and clinical factors to predict the therapeutic dose of warfarin. *Clin Pharmacol Ther* 2008;84:326–31.
- Furie B. Do Pharmacogenetics Have a Role in the Dosing of Vitamin K Antagonists? *N Engl J Med* 2013;369:2345–46.
- Davey Smith G, Egger M. Incommunicable Knowledge? Interpreting and Applying the Results of Clinical Trials and Meta-Analyses. *J Clin Epidemiol* 1998;51:289–95.
- Poisson SD, Dulong PL, Larrey DJ *et al*. Statistical research on conditions caused by calculi by Doctor Civiale. *Int J Epidemiol* 2001;30:1246–49.
- Rosser Matthews J. Commentary: The Paris Academy of Science report on Jean Civiale’s statistical research and the 19th century background to evidence-based medicine. *Int J Epidemiol* 2001;30:1249–50.
- Black N. Commentary: That was then, this is now. *Int J Epidemiol* 2001;30:1251.
- Tröhler U. Commentary: ‘Medical art’ versus ‘medical science’: J. Civiale’s statistical research on conditions caused by calculi at the Paris Academy of Sciences in 1835. *Int J Epidemiol* 2001;30:1252–53.
- Vandenbroucke JP. Commentary: Treatment of bladder stones and probabilistic reasoning in medicine: an 1835 account and its lessons for the present. *Int J Epidemiol* 2001;30:1253–58.
- Davey Smith G, Ebrahim S. ‘Mendelian randomization’: can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;32:1–22.
- Davey Smith G, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *Int J Epidemiol* 2004;33:30–42.

41. Davey Smith G, Lawlor DA, Harbord R, Timpson NJ, Day I, Ebrahim S. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Med* 2007;4:e352.
42. Davey Smith G. Random Allocation in Observational Data: How small but robust effects could facilitate hypothesis-free causal inference. *Epidemiology* 2011;22:460–63.
43. Mendelson MM, Marioni RE, Joehanes R *et al*. Association of Body Mass Index with DNA Methylation and Gene Expression in Blood Cells and Relations to Cardiometabolic Disease: A Mendelian Randomization Approach. *PLoS Med* 2017;14:e1002215.
44. Ronksley PE, Brien SE, Turner BJ, Mukamal KJ, Ghali WA. Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis. *BMJ* 2011;342:d671.
45. Mukamal KJ, Mackey RH, Kuller LH *et al*. Alcohol consumption and lipoprotein subclasses in older adults. *J Clin Endocrinol Metab* 2007;92:2559–66.
46. Greenfield JR, Samaras K, Jenkins AB, Kelly PJ, Spector TD, Campbell LV. Moderate alcohol consumption, estrogen replacement therapy, and physical activity are associated with increased insulin sensitivity: is abdominal adiposity the mediator? *Diabetes Care* 2003;26:2734–40.
47. Imhof A, Froehlich M, Brenner H, Boeing H, Pepys MB, Koenig W. Effect of alcohol consumption on systemic markers of inflammation. *Lancet* 2001;357:763–67.
48. Albert MA, Glynn RJ, Ridker PM. Alcohol consumption and plasma concentration of C-reactive protein. *Circulation* 2003;107:443–47.
49. Mukamal KJ, Jadhav PP, D'Agostino RB, *et al*. Alcohol consumption and herostatic factors: analysis of the Framingham Offspring Cohort. *Circulation* 2001;104:1367–73.
50. Di Castelnuovo A, Costanzo S, Bagnardi V, Donati MB, Iacoviello L, de Gaetano G. Alcohol dosing and total mortality in men and women: an updated meta-analysis of 34 prospective studies. *Arch Intern Med* 2006;166:2437–45.
51. Marmot MG. Alcohol and coronary heart disease. *Int J Epidemiol* 2001;30:72–79.
52. Holmes MV, Dale CE, Zuccolo L *et al*. Association between alcohol and cardiovascular disease: Mendelian randomization analysis based on individual participant data. *BMJ* 2014;349:g4164.
53. Silverwood RJ, Holmes MV, Dale CE *et al*; on behalf of the Alcohol-ADH1B Consortium. Testing for non-linear causal effects using a binary genotype in a Mendelian randomization study: application to alcohol and cardiovascular traits. *Int J Epidemiol* 2014;43:1781–90.
54. Alcohol Policy Team, Department of Health. *How to Keep Health Risks From Drinking Alcohol to a Low Level*. Government response to the public consultation. London: Department of Health, 2016.
55. Kelly YJ, Sacker A, Gray R *et al*. Light drinking during pregnancy: still no increased risk for socioemotional difficulties or cognitive deficits at 5 years of age? *J Epidemiol Community Health* 2012;66:41–48.
56. Zuccolo L, Lewis SJ, Davey Smith G *et al*. Prenatal alcohol exposure and offspring cognition and school performance. A 'Mendelian randomization' natural experiment. *Int J Epidemiol* 2013;42:1358–70.
57. Oh J, Teoh H, Leiter LA. Should C-Reactive Protein Be a Target of Therapy? *Diabetes Care* 2011; 34 (Suppl 2):S155–60.
58. C Reactive Protein Coronary Heart Disease Genetics Collaboration (CCGC). Association between C reactive protein and coronary heart disease: Mendelian randomisation analysis based on individual participant data. *BMJ* 2011;342:d548.
59. Barter PJ, Caulfield M, Eriksson M *et al*. Effects of torcetrapib in patients at high risk for coronary events. *N Engl J Med* 2007;357:2109–22.
60. Sofat R, Hingorani AD, Smeeth L *et al*. Separating the mechanism-based and off-target actions of CETP-inhibitors using CETP gene polymorphisms. *Circulation* 2010;121:52–62.
61. Voight BF, Peloso GM, Orho-Melander M *et al*. Plasma HDL cholesterol and risk of myocardial infarction: a Mendelian randomisation study. *Lancet* 2012;380:572–80.
62. Kosmas CE, DeJesus E, Rosario D, Vittorio TJ. CETP Inhibition: Past Failures and Future Hopes. *Clin Med Insights Cardiol* 2016;10:37–42.
63. HPS3/TIMI 55–REVEAL. Randomised Evaluation of the Effects of Anacetrapib through Lipid-modification. <https://www.ctsu.ox.ac.uk/research/hps3-timi55-reveal> (4 January 2017, date last accessed).
64. Swerdlow DI, Preiss D, Kuchenbaecker KB *et al*. HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials. *Lancet* 2015;385:351–61.
65. Everett BM, Smith RJ, Hiatt WR. Reducing LDL with PCSK9 Inhibitors - The Clinical Benefit of Lipid Drugs. *N Engl J Med* 2015;373:1588–91.
66. Ference B, Robinson J, Brook R *et al*. Variation in PCSK9 and HMGCR and Risk of Cardiovascular Disease and Diabetes. *N Engl J Med* 2016;375:2144–50.
67. Mokry LE, Ahmad O, Forgetta V, Thanassoulis G, Richards JB. Mendelian randomisation applied to drug development in cardiovascular disease: a review. *J Med Genet* 2015;52:71–79.
68. Martin J, Tilling K, Hubbard L *et al*. Association of Genetic Risk for Schizophrenia With Nonparticipation Over Time in a Population-Based Cohort Study. *Am J Epidemiol* 2016;183: 1149–58.
69. Rothman KJ, Gallacher JEJ, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol* 2013;42:1012–14.
70. Keiding N, Louis TA. Perils and potentials of self-selected entry to epidemiological studies and surveys. *J R Stat Soc* 2016; 179(Part 2):319–76.
71. Richiardi L, Pizzi C, Pearce N. Commentary. Representativeness is usually not necessary and often should be avoided. *Int J Epidemiol* 2013;42:1018–22.
72. Cole SR, Platt RW, Schisterman EF *et al*. Illustrating bias due to conditioning on a collider. *Int J Epidemiol* 2010;39:417–20.
73. Ebrahim S, Davey Smith G. Commentary: Should we always deliberately be non-representative? *Int J Epidemiol* 2013;42: 1022–26.
74. Rothman KJ, Gallacher JEJ, Hatch EE. Rebuttal: When it comes to scientific inference, sometimes a cigar is just a cigar. *Int J Epidemiol* 2013;42:1026–28.
75. Sterne J, Davey Smith G. Sifting the evidence—what's wrong with significance tests? *BMJ* 2001;322:226.

76. Lawlor DA, Davey Smith G, Ebrahim S. Socioeconomic Position and Hormone Replacement Therapy Use: Explaining the Discrepancy in Evidence From Observational and Randomized Controlled Trials. *Am J Public Health* 2004;94:2149–54.
77. Lawlor DA, Davey Smith G, Bruckdorfer KR, Kundu D, Ebrahim S. Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? *Lancet* 2004;363:1724–27.
78. Lawlor DA, Ebrahim S, Kundu D, Bruckdorfer KR, Whincup PH, Davey Smith G. Vitamin C is not associated with coronary heart disease risk once life course socioeconomic position is taken into account: prospective findings from the British Women's Heart and Health Study. *Heart* 2005;91:1086–87.
79. He K, Merchant A, Rimm EB *et al.* Folate, vitamin B6, and B12 intakes in relation to risk of stroke among men. *Stroke* 2004;35:169–74.
80. Al-Delaimy WK, Rexrode KM, Hu FB *et al.* Folate intake and risk of stroke among women. *Stroke* 2004;35:1259–63.
81. Hak AE, Ma J, Powell CB *et al.* Prospective study of plasma carotenoids and tocopherols in relation to risk of ischemic stroke. *Stroke* 2004;35:1584–88.
82. Hak AE, Stampfer MJ, Campos H *et al.* Plasma carotenoids and tocopherols and risk of myocardial infarction in a low-risk population of US male physicians. *Circulation* 2003;108:802–07.
83. Ascherio A, Rimm EB, Hernán MA *et al.* Relation of consumption of vitamin E, vitamin C, and carotenoids to risk for stroke among men in the United States. *Ann Intern Med* 1999;130:963–70.
84. Pigou AC. Alcoholism and heredity. *Westminster Gazette*, 2nd February 1911. Reprinted in the *Int J Epidemiol* 2017, in press.
85. Munafo MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider Scope: How selection bias can induce spurious associations. *bioRxiv preprint first posted online Oct. 7, 2016.* <http://dx.doi.org/10.1101/079707>
86. Evans DM, Visscher PM, Wray NR. Harnessing the information contained within genome wide association studies to improve individual prediction of complex disease risk. *Hum Mol Genet* 2009;18:3525–31.
87. Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* 2012;28:2540–42.
88. Bulik-Sullivan B, Finucane HK, Anttila V *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* 2015;47:1236–41.
89. Day FR, Loh P-R, Scott RA, Ong KK, Perry JRB. A Robust Example of Collider Bias in a Genetic Association Study. *Am J Hum Genet* 2016;98:392–93.
90. Yaghootkar H, Bancks M, Jones S *et al.* Quantifying the extent to which index event biases influence large genetic association studies. *Hum Mol Genet* 2016, December 30. doi:10.1093/hmg/ddw433.
91. Aschard H., Vilhjálmsson BJ, Joshi AD, *et al.* Adjusting for heritable covariates can bias effect estimates in genome-wide association studies *The American Journal of Human Genetics* 2015;96:329–339.
92. Jones HJ, Stergiakouli E, Tansey KE *et al.* Phenotypic Manifestation of Genetic Risk for Schizophrenia During Adolescence in the General Population. *JAMA Psychiatry* 2016;73:221–28.
93. Greenwood M. *Epidemics and Crowd Diseases. An Introduction to the Study of Epidemiology.* London: Williams & Norgate, 1935.
94. Rose G. *The Strategy of Preventive Medicine.* Oxford, UK: Oxford University Press, 1992.
95. Susser M. Epidemiology in the United States after World War II. The evolution of technique. *Epidemiol Rev* 1985;7:147–77.
96. Stallones RA. To advance epidemiology. *Ann Rev Public Health* 1980;1:69–82.
97. Morris JN. Modern epidemiology? *J Epidemiol Community Health* 1988;42:100.
98. Susser M, Susser E. Choosing a future for epidemiology. II. From black box to Chinese boxes and eco-epidemiology. *Am J Public Health* 1996;86:647–77.
99. Vandembroucke JP, Broadbent A, Pearce N. Causality and causal inference in epidemiology: the need for a pluralistic approach. *Int J Epidemiol* 2016;45:1776–86.
100. Pearce N, Vandembroucke JP. Causation, mediation, and explanation. *Int J Epidemiol* 2016;45:1915–22.
101. Krieger N, Davey Smith G. The tale wagged by the DAG: broadening the scope of causal inference and explanation for epidemiology. *Int J Epidemiol* 2016;45:1787–808.
102. Weed DL. Causal inference in epidemiology: potential outcomes, pluralism and peer review. *Int J Epidemiol* 2016;45:1838–40.
103. Chiolerio A. Letter to the Editor: Counterfactual and interventionist approach to cure risk factor epidemiology. *Int J Epidemiol* 2016;45:2202–03.
104. Lawlor DA, Davey Smith G, Tilling K. Triangulation in aetiological epidemiology. *Int J Epidemiol* 2016;45:1866–86.
105. VanderWeele TJ, Hernán MA, Tchetgen Tchetgen EJ, Robins JM. Letter to the Editor: Re: Causality and causal inference in epidemiology: the need for a pluralistic approach. *Int J Epidemiol* 2016;45:2199–200.
106. VanderWeele TJ. On causes, causal inference, and potential outcomes. *Int J Epidemiol* 2016;45:1809–16.
107. Blakely T, Lynch J, Bentley R. DAGs and the restricted potential outcomes approach are tools, not theories of causation. *Int J Epidemiol* 2016;45:1835–37.
108. Daniel RM, De Stavola BL, Vansteelandt S. The formal approach to quantitative causal inference in epidemiology: misguided or misrepresented? *Int J Epidemiol* 2016;45:1817–29.
109. Robins JM, Weissman MB. Counterfactual causation and streetlamps. What is to be done? *Int J Epidemiol* 2016;45:1830–35.
110. Kreiger N, Davey Smith G. FACEing reality: productive tensions between our epidemiological questions, methods, and mission. *Int J Epidemiol* 2016;45:1852–65.
111. Broadbent A, Vandembroucke JP, Pearce N. Letter to the Editor: Authors' reply to: VanderWeele *et al.*, Chiolerio, and Schooling *et al.* *Int J Epidemiol* 2016;45:2203–05.
112. Textor J, van der Zander B, Gilthorpe MS, LiSkiewicz M, Ellison GTH. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. *Int J Epidemiol* 2016;45:1887–94.

113. Byberg KK, Eide GE, Forman MR, Júlíusson PB, Øymar K. Body mass index and physical activity in early childhood are associated with atopic sensitization, atopic dermatitis and asthma in later childhood. *Clin Transl Allergy* 2016;6:33.
114. Wadonda-Kabondo N, Sterne JAC, Golding J, Kennedy CTC, Archer CB, Dunnill MGS. Association of parental eczema, hayfever, and asthma with atopic dermatitis in infancy: birth cohort study. *Arch Dis Childhood* 2004;89:917–21.
115. Chen JT, Krieger N, Van Den Eeden SK *et al*. Different slopes for different folks: socioeconomic and racial/ethnic disparities in asthma and hay fever among 173,859 U.S. men and women. *Environ Health Perspect* 2002;110(Suppl 2):211–16.
116. Forastiere F, Agabiti N, Corbo GM *et al*. Socioeconomic status, number of siblings, and respiratory infections in early life as determinants of atopy in children. *Epidemiology* 1997;85: 566–570.
117. Uphoff E, Cabieses B, Pinart M *et al*. A systematic review of socioeconomic position in relation to asthma and allergic diseases. *Eur Respir J* 2015;46:364–74.
118. Galobardes B, Granell R, Sterne J *et al*. Childhood Wheezing, Asthma, Allergy, Atopy, and Lung Function: Different Socioeconomic Patterns for Different Phenotypes. *Am J Epidemiol* 2015;182:763–74.
119. Davey Smith G. The Wright Stuff. Genes in the interrogation of correlation and causation. *Eur J Pers* 2012;26:391–413.
120. Mukherjee S. *The Laws of Medicine*. London: Simon Shuster, 2015.
121. Lipton P. Inference to the best explanation. In: Newton-Smith WH (ed). *A Companion to the Philosophy of Science*. New York, NY: Blackwell, 2000.
122. Keyes KM, Galea S. *Population Health Sciences*. Oxford, UK: Oxford University Press, 2016.
123. World Health Organisation. Health topics Epidemiology. <http://www.who.int/topics/epidemiology/en/> WHO 2017 (25 January 2017, date last accessed).
124. Press G. *12 Big Data Definitions: What's Yours?* 2014. <http://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours/#259cfe2521a9> (25 January 2017, date last accessed).
125. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst*. 2014;2:3.
126. Herrett E, Gallagher AM, Bhaskaran K *et al*. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015;44:827–36.
127. Smeeth L, Cook C, Fombonne E *et al*. MMR vaccination and pervasive developmental disorders: a case-control study. *Lancet* 2004;364:963–69.
128. Freemantle N, Ray D, McNulty D *et al*. Increased mortality associated with weekend hospital admission: a case for expanded seven day services? *BMJ* 2015;351:h4596.
129. Freemantle N, Pagano D. Electronic response to Freemantle N, *et al*. Increased mortality associated with weekend hospital admission: a case for expanded seven day services? *BMJ* 2015;351:h4596.
130. Goldacre B. All BMJ research papers should share their analytic code. *BMJ* 2016;352:i886.
131. Achinger SG, Go AS, Ayus JC. Getting it straight: avoiding blunders while criticizing a peer's work. *Int J Epidemiol* 2016;45:619–20.
132. Hanley JA. Immortal time bias. Response to: Achinger, Go and Ayus. *Int J Epidemiol* 2016;45:965–66.
133. <http://www.evidenceaction.org/dewormtheworld/> (Last accessed 12 October 2015).
134. Miguel E, Kremer M. Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica* 2004;72:159–217.
135. Bleakley H. Disease and development: evidence from hookworm eradication in the American South. *Q J Econ* 2007;122: 73–117.
136. Baird S, Hicks JH, Miguel E, Kremer M. *Worms at Work: Long-run Impacts of Child Health Gains*. http://eml.berkeley.edu/~emiguel/pdfs/miguel_wormsatwork.pdf (12 January 2017, date last accessed).
137. Aiken ADC, Davey C, Hargreaves J, Hayes R. Re-analysis of health and educational impacts of a school-based deworming programme in western Kenya: a pure replication. *Int J Epidemiol* 2015;44:1572–80.
138. Davey C, Aiken A, Hayes R, Hargreaves J. Re-analysis of health and educational impacts of a school-based deworming programme in Western Kenya: a statistical replication of a cluster quasi-randomized stepped-wedge trial. *Int J Epidemiol* 2015;44:1581–92.
139. Garner P, Taylor-Robinson D, Sachdev HS. Commentary: Replication of influential trial helps international policy. *Int J Epidemiol* 2015;44:1599–601.
140. Goldacre B. *Scientists Are Hoarding Data and It's Ruining Medical Research*. <http://www.buzzfeed.com/bengoldacre/deworming-trials> (25 January 2017, date last accessed).
141. Taylor-Robinson DC, Maayan N, Soares-Weiser K, Donegan S, Garner P. Deworming drugs for soil-transmitted intestinal worms in children: effects on nutritional indicators, haemoglobin and school performance. *Cochrane Database Syst Rev* 2015;CD000371.
142. Awasthi S, Peto R, Read S *et al*. Population deworming every 6 months with albendazole in 1 million pre-school children in North India: DEVTA, a cluster-randomised trial. *Lancet* 2013;381:1478–86.
143. Jullien S, Sinclair D, Garner P. The impact of mass deworming programmes on schooling and economic development: an appraisal of long-term studies. *Int J Epidemiol* 2016;45:2140–53.
144. Ahuja A, Baird S, Hamory Hicks J, Kremer M, Miguel E, Powers S. When Should Governments Subsidize Health? The Case of Mass Deworming. *World Bank Econ Rev* 2015;29(Suppl1): S9–24.
145. World Health Organisation. Neglected tropical diseases. WHO advisory body upholds continued deworming of children against soil – transmitted helminthiases. Statement from the WHO Strategic and Technical Advisory Group for Neglected Tropical Diseases. http://www.who.int/neglected_diseases/news/update-deworming-children/en/ (12 October 2016, date last accessed).
146. MacAskill W. *Doing Good Better: Effective Altruism and a Radical New Way to Make a Difference*. London: Guardian Faber, 2015.
147. Cochrane Library. *The Deworming Debate*. 2015. <http://www.cochrane.org/news/deworming-debate> (25 January 2017, date last accessed).
148. Baird S, Hamory Hicks J, Kremer M, Miguel E. Commentary: Assessing long-run deworming impacts on education and economic outcomes: a comment on Jullien, Sinclair and Garner. *Int J Epidemiol* 2016;45:2153–56.

149. Ozier O. Commentary: Fixed effects and risks of miscommunication: a comment on Jullien, Sinclair and Garner. *Int J Epidemiol* 2016;45:2156–58.
150. Croke K. Commentary: Exploiting randomized exposure to early childhood deworming programmes to study long-run effects: A research programme in progress. *Int J Epidemiol* 2016;45:2159–61.
151. Hillebrandt H. Commentary: Three ways to falsify the case for mass deworming against soil-transmitted helminths. *Int J Epidemiol* 2016;45:2168–70.
152. Humphreys M. Commentary: Biases in the assessment of long-run effects of deworming. *Int J Epidemiol* 2016;45:2163–65.
153. Roodman D. Commentary: From conclusions to decisions: a comment on Jullien et al. *Int J Epidemiol* 2016;45:2166–68.
154. Aiken A, Wood BDK. Commentary: The tide continues to move on mass deworming—where are we now? *Int J Epidemiol* 2016;45:2161–63.
155. Wallach JD, Basu S. Commentary: The challenge of evaluating and improving evidence when research lacks reproducibility. *Int J Epidemiol* 2016;45:2171–72.
156. Black, D. *Inequalities in Health*. London: Department of Health and Social Security, 1980.
157. Capewell S, Graham H. Will Cardiovascular Disease Prevention Widen Health Inequalities? *PLoS Med* 2010;7:e1000320.
158. Lorenc T, Petticrew M, Welch V, Tugwell P. What types of interventions generate inequalities? Evidence from systematic reviews. *J Epidemiol Community Health* 2013;67:190–93.
159. UCL Institute of Health Equity. *The Marmot Review – Fair Society Healthy Lives*. <http://www.instituteofhealthequity.org/presentations/presentation-slides> (15 December 2016, date last accessed).
160. Marmot M. *The Health Gap. The Challenge of an Unequal World*. London: Bloomsbury, 2015.
161. Marmot M, Bell R. Social inequalities in health: a proper concern of epidemiology. *Ann Epidemiol* 2016;26:238–40.
162. McCartney D, Scarborough P, Webster P, Rayner M. Trends in social inequalities for premature coronary heart disease mortality in Great Britain, 1994–2008: a time trend ecological study. *BMJ Open* 2012;2:e000737.
163. Dorling D. *Inequality and the 1%*. London: Verso, 2014.
164. Jones O. *The Establishment. And How They Get Away With It*. London: Penguin, 2015.
165. Chittleborough CR, Mittinty MN, Lawlor DA, Lynch JW. Effects of simulated interventions to improve school entry academic skills on socioeconomic inequalities in educational achievement. *Child Dev* 2014;85:2247–62.
166. Dorling D. The Scottish mortality crisis. *The Geographer* 2016;Summer:8.
167. Ebrahim S. Surveillance and monitoring: a vital investment for the changing burdens of disease. *Int J Epidemiol* 2011;40:1139–43.
168. Farzadfar F, Finucane MM, Danaei G *et al.*; on behalf of the Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Cholesterol). National, regional, and global trends in serum total cholesterol since 1980: systematic analysis of health examination surveys and epidemiological studies with 321 country-years and 30 million participants. *Lancet* 2011;377:578–86.
169. Danaei G, Finucane MM, Lu Y, *et al.*; on behalf of the Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Blood Glucose). National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 27 million participants. *Lancet* 2011;378:31–40.
170. Byass P. Cause-specific mortality findings from the Global Burden of Disease project and the INDEPTH Network. *Lancet Glob Health* 2016;4:e785–86.
171. Sankoh O, Byass P. The INDEPTH Network: filling vital gaps in global epidemiology. *Int J Epidemiol* 2012;41:579–88.
172. Lopez A. Global Health Targets for 2015. *Int J Epidemiol* 2015;44:9–11.
173. United Nations Synthesis Report of the Secretary-General on the Post-2015 Agenda. *The Road to Dignity by 2030: Ending Poverty, Transforming All Lives and Protecting the Planet*. New York, NY: United Nations Organization, 2014.
174. GBD 2015 SDG Collaborators. Measuring the health-related Sustainable Development Goals in 188 countries: a baseline analysis from the Global Burden of Disease Study 2015. *Lancet* 2016;388:1813–50.
175. Ferrie JE. Incidents, incidence and golden eggs. *Int J Epidemiol* 2012;41:329–32.
176. Ferrie JE. Data resources from Case No. 28 to datasets of millions. *Int J Epidemiol* 2012;41:1495–500.
177. Kelaher M, Rahadi A, Ng, L, Knight, K. Equity in global health research in the new millennium: trends in first authorship among low- and middle-income country researchers 1990–2013. *Int J Epidemiol* 2016;45:2174–83.
178. Långström N, Babchishin KM, Fazel S, Lichtenstein P, Frisell T. Sexual offending runs in families: a 37-year nationwide study. *Int J Epidemiol* 2015;44:713–20.
179. Mazières D, Kohler E. Get me off your fucking mailing list. *Ninth World Multi-Conference on Systemics, Cybernetics, and Informatics, Orlando, FL, July 2005*. <http://www.scs.stanford.edu/~dm/home/papers/remove.pdf> (16 March 2017, date last accessed).
180. *International Journal of Advanced Computer Technology*. https://en.wikipedia.org/wiki/International_Journal_of_Advanced_Computer_Technology (18 December 2016, date last accessed).
181. Wellcome Open Research. <https://wellcomeopenresearch.org/> ISSN 2398-502X (5 December 2016, date last accessed).
182. F1000 Research. *Open for Science*. <https://f1000research.com/> (5 December 2016, date last accessed).
183. Miller M. Hunter S. Thompson's Rejection Letter to a Famous Author: 'Cheapjack Scum': The man had a way with insults. <http://www.esquire.com/entertainment/books/news/a41028/hunter-s-thompson-rejection-letter-anthony-burgess/> (15 December 2016, date last accessed).
184. Larkin P. High Windows. In: *Collected Poems*. London: Faber and Faber; 2001. Available from <https://www.poetryfoundation.org/poems-and-poets/poems/detail/48417> (16 March 2017, date last accessed).