Editorial

The future of epidemiology: methods or matter?

Shah Ebrahim¹, Jane E Ferrie² and George Davey Smith³*

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 Lewontin²¹,²² 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 drinkers45–49 — 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. 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’. 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. This and other new evidence appears to have influenced alcohol guidance in the UK to promote lower levels of consumption. 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. 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, a paper in the *IJE* which applied Mendelian randomization to the question reached the conclusion that any alcohol consumption was detrimental, 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. Mendelian randomization studies have now shown convincingly that CRP is not causally related to cardiovascular disease, 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. 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. However, Mendelian randomization studies indicate that genetic variants that raise HDL cholesterol are not associated with lower risk of cardiovascular disease, predictions borne out by clinical trials of CETP inhibitors mounted by Eli Lilly, Roche and Pfizer. Merck has continued to support its trial, expected to report in 2017, of a CETP-

![Figure 1](image-url). Behavioural difficulties in offspring and the behavioural and socioeconomic characteristics of their mothers, by drinking category. Figure based on data provided in reference 55.
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). Trials of a new class of cholesterol-lowering drug, proprotein convertase subtilisin-kexin type 9 (PCSK9) inhibitors 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. 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.

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. 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. 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’. 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. 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. 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 bias – 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.72 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, 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.86 This, in turn, can potentially lead to bias. Studies using polygenic scores, genome-wide allele 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). Less contrived examples of this and related problems are beginning to appear. 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. 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.

Causal inference and DAGs

Despite eloquent presentations of the purpose and uses of epidemiology over the past century, concerns about its direction have been voiced repeatedly. 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.

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. 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’. However, the DAG does not provide a comprehensive picture. For example, it does not include
population 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 reinterpretation 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. 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. The investigators refused to share their analytic code on grounds of research governance, 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. 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.

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. 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, augmented by data from a study from 1910, and an unpublished report of a follow-up of the original Kenya study. 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. The re-analyses found a small effect of deworming on school attendance but no clear effect on examination performance. 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.

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, and a mega-trial of two million children in India failed to demonstrate benefits from deworming on mortality or weight gain. 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, 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.’

Are downstream benefits, such as economic development, feasible in light of no clear benefits for upstream outcomes? Despite the growing evidence showing that deworming has little to offer at a population level, it remains popular with the World Health Organization and the ‘Effective Altruism’ movement, who have claimed that deworming is more effective than providing textbooks or teachers in Kenya. 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. The commentators comprise the authors of the original working papers, GiveWell (an organization that assesses evidence on interventions supported by charities) and other scientists with methodological interests. 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. 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 Report 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. 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. 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).

By contrast, when the relative rates of decline are examined (Figure 6) 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?’, in order to develop a better understanding of how those in power maintain their grip on society. There is a role for epidemiology here...
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, 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 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). Statistical models filled in the gaps by ‘borrowing’ from other countries in the South Asia region, modelling
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
involvement of LMIC researchers. A variation we observed was token authorship for investigators from LMICs in the middle of the author list – the ‘thinely 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
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 2001 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 or very little editorial oversight or input, or appear to be compiled by machines (see Figure 9). 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’ initiative 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. 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...’ (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). We are no longer young, and don’t deserve such.

Acknowledgments

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

References

36. Black N. Commentary: That was then, this is now. Int J Epidemiol 2001;30:1251.


149. Ozier O. Commentary: Fixed effects and risks of miscommunica-
tion; a comment on Jullien, Sinclair and Garner. Int J Epidemiol
2012;41:329–32.
150. Ferrie JE. Data resources from Case No. 28 to datasets of mil-
151. 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-
2015;44:9–11.
153. 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
154. 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
155. Ferrie JE. Incidents, incidence and golden eggs. Int J Epidemiol
2012;41:329–32.
157. Aiken A, Wood BDK. Commentary: The tide continues to
move on mass deworming–where are we now? Int J Epidemiol
2016;45:2161–63.
158. Black, D. Inequalities in Health. London: Department of
159. 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
160. Ebrahim S. Surveillance and monitoring: a vital investment for the
161. Croke K. Commentary: Exploiting randomized exposure to early
childhood deworming programmes to study long-run effects:
A research programme in progress. Int J Epidemiol
162. Dorling D. The Scottish mortality crisis. The Geographer
2016;Summer:8.
163. Chittleborough CR, Mittinty MN, Lawlor DA, Lynch JW. Effe-
ts of simulated interventions to improve school entry aca-
demic skills on socioeconomic inequalities in educational
165. Fairuz F, Finucane MM, Danaei G et al.; on behalf of the Global
Burden of Metabolic Risk Factors of Chronic Diseases Col-
laborating 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.
166. Luo N, Qureshi A, Kulldorff M et al.; on behalf of the Global
Burden of Metabolic Risk Factors of Chronic Diseases Col-
laborating 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 mil-