

UNIVERSITY OF COPENHAGEN



An Elusive Evidence Base: The Construction and Governance of Randomized Controlled Trials

Wahlberg, Ayo; McGoey, Linsey

Published in:
BioSocieties

DOI:
[10.1017/S1745855207005017](https://doi.org/10.1017/S1745855207005017)

Publication date:
2007

Citation for published version (APA):
Wahlberg, A., & McGoey, L. (2007). An Elusive Evidence Base: The Construction and Governance of Randomized Controlled Trials. *BioSocieties*, 2(1), 1-10. <https://doi.org/10.1017/S1745855207005017>

An Elusive Evidence Base: The Construction and Governance of Randomized Controlled Trials

Ayo Wahlberg and Linsey McGoey

BIOS, LSE, Houghton Street, London WC2A 2AE, UK
E-mail: a.j.wahlberg@lse.ac.uk

In 1987, just a year before his death, the British epidemiologist Archie Cochrane gave one of his last public interviews. Cochrane had led an illustrious career, helping to establish the fields of epidemiology and public health in Britain and internationally. It was, however, the publication of a slim monograph, *Effectiveness and efficiency: Random reflections on health services* (1972), that sealed his reputation. Here he called for an international register of randomized controlled trials (RCTs) that would allow researchers the ability to assess the efficacy of health interventions. The call to arms had a strong influence on the next generation of clinical epidemiologists and clinicians, including the McMaster team that coined the term evidence-based medicine (EBM) in the early 1990s (Mykhalovskiy and Weir 2004: 1059). Twenty years later, it appears Cochrane's dream of establishing public databases of clinical trials has been largely realized, while his call for the development of open and transparent criteria for appraising the validity of published medical research has become fundamental to an international EBM movement. Today, governments in North America, Britain, Western Europe and Australia fund institutions (such as the UK's National Institute for Health and Clinical Excellence, or NICE) which commission research, collate evidence and produce evidence-based guidelines that physicians are encouraged to use in their clinical practice.

Nevertheless, early intimations of the now widespread criticisms of EBM in our culture were actually visible from the very start. In the 1987 interview, in which he mused on his achievements over the years, Cochrane recalled:

... the use of randomised clinical trials has increased. The only tragedy is that it hasn't spread throughout the world. Very few are done in France and Germany and none in Russia and it's really confined to Canada, the United States, Scandinavia and United

Ayo Wahlberg is a Research Fellow at the BIOS Centre, LSE, working as part of BIONET, an EU–China collaboration on the ethical governance of biomedical research. He holds a PhD in Sociology from the LSE and an MSc in International Development Studies and Social Science from Roskilde University, Denmark. His doctoral work was on contemporary controversies surrounding quackery, efficacy and life in Vietnamese and British herbal medicine. Publications include 'Bio-politics and the promotion of traditional herbal medicine in Vietnam' (in *Health: an Interdisciplinary Journal for the Social Study of Health, Illness and Medicine*, 10, 123–147).

Linsey McGoey is a PhD Candidate at the BIOS Centre, LSE, working on a thesis entitled 'The moral authority of objectivity in medicine'. With Scott Vrecko, she is co-founder of the Neuroscience and Society Network, an international network of social science and life science researchers investigating various aspects of the contemporary brain sciences. Her article 'On the will to ignorance in bureaucracy' is forthcoming in the journal *Economy and Society* 36(2).

Kingdom. If all the countries would join in we could get the answers far more rapidly. The attitude of the Germans is quite extraordinary. I was over there only about three years ago and my German was still very good and I gave a standard lecture on the use of the randomised controlled trials in organising the treatment of heart disease. A very straightforward thing: the randomised controlled trials of place of treatment, length of stay, and all that sort of thing, and I've never had a more riotous reception. My god they were rude to me, in German. Fortunately I can fight back in German and we had a real battle. They considered me utterly unethical and I complained that they were unethical: they weren't checking that they were giving their patients the best treatment. It went on for nearly three-quarters of an hour. I got pretty tired. The real story comes at the end. The professor took me out and gave me dinner and on the way out in excellent English said, 'You know, Dr. Cochrane, you don't seem to understand about controlled trials. Controlled trials are done by the pharmaceutical industry. Gentlemen don't do them.' (Cochrane, 1987: 14)

For decades, in fact, it has often been gentlemen scientists and clinicians themselves who have been the most vocal in raising internal critiques of some of the problems facing evidence-based medicine. Practitioners have been at the forefront of efforts to identify methodological obstacles that can distort the validity of trial results, calling attention to problems of, among other things, patient inclusion criteria, the recording of trial endpoints, and the influence of selective trial publication on the development of evidence-based guidelines (Armstrong, 2002; Chalmers, 1990; Healy, 2001; Jørgensen *et al.*, 2006; Moncrieff, 2004).

It is only more recently that sociologists, anthropologists and ethicists have joined practitioners in scrutinizing the influence of randomized controlled trials and their use within evidence-based medicine. Clinicians themselves have contributed so extensively to internal critiques of the rise and practice of evidence-based medicine that the more recent sociological attention has struck some practitioners as at best redundant, and at worst arrogant—arrogant for approaching the field as though methodological problems surrounding the construction of RCTs had not themselves been identified by practitioners.

With this special issue of *BioSocieties*, it has been our aim to encourage empirical investigation of, as well as theoretical incursions into, the political, regulatory and commercial contexts in which the model of evidence-based medicine has grown to be an increasingly dominant paradigm across health-care systems worldwide.¹ As Adriana Petryna writes in this issue:

Much has been written about the social construction of the evidentiary process in clinical trials—how the evidence-making process is itself tainted. But not much has been written about how scientific models affect political processes downstream.

It is in this sense that the issue's focus on the construction and governance of randomized controlled trials should be approached. While authors certainly do address many of the

¹ The idea for this issue emerged from a two-day symposium held at the London School of Economics on 8–9 June 2006. 'Searching for gold standards—The construction and governance of RCTs and EBM in psychiatry', was jointly organized by the Institute of Psychiatry, King's College London and the BIOS Centre. Our thanks to Derek Bolton, Nikolas Rose and George Szumkler for their oversight and support of the symposium.

controversies that have surrounded the evidence-making process in medicine, they have done so with the aim of illuminating some of the very practical and technical concerns that surround current strategies of clinical testing. There are many important ethical and moral questions to be asked of these strategies, but these questions do not arise from some kind of *ex post facto* revelation that evidence bases were ‘after all’ constructed. Instead, they are partially determined by the practices of construction and governance that operate in randomized controlled trials. Realizing this allows us then to ask, as Nikolas Rose has done, ‘Where do [forms of knowledge/evidence] emerge? Which are the authorities who are able to pronounce upon them? Through what concepts and explanatory regimes are they investigated?’ (1999: x).

What’s in a name? Debates over terminology

No proper analysis of EBM can proceed in the absence of a working definition of the concept, and a clarification of one’s choice of terminology.² There are just too many heterogeneous and often conflicting assemblies of clinicians, policy-makers, regulators, corporate parties and other stakeholders involved in the industry to assume easy consensus. On the one hand, Sackett’s oft-cited definition of evidence-based medicine as the ‘the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients’ (1996: 71) highlights its role as an aid in clinical judgement. Doctors have long endeavoured to keep themselves up to date with the latest medical advances and findings, not to mention drawing from their own clinical experience, in order to optimize their decisions about treatment for individual patients.

Yet, with so much new trial data emerging and in increasingly specialized fields, it is often pointed out that an individual doctor’s efforts to keep him or herself *à jour* has become increasingly difficult. This has led some to define EBM as a strategy of rationalization whereby certain individuals are delegated the task of keeping abreast of and collating latest advances, trial evidence and observational studies. It is through this kind of work that clinical guidelines can be drawn up according to Cochrane-inspired hierarchies of evidence, as a means of providing doctors with guidance on how to proceed with decisions about treatment in individual cases.

Yet, as Cochrane was at pains to draw out in his 1972 monograph, contributing to the collecting and collating of ‘current best evidence’ was not only about ensuring that *individual* patients were given the most effective therapies; it was equally about ensuring that *national* health delivery systems worked as efficiently as possible. As growing proportions of national budgets come to be spent on health care, especially in industrialized countries, cost-conscious, efficient spending becomes an imperative: ‘Decisions taken at higher levels are so dependent on accurate information about the effectiveness and efficiency of the treatment of individual diseases’ (1972: 31). Treatment guidelines prepared by organizations such as NICE in the UK and Health Maintenance Organizations (HMOs) in the US are informed not only by ‘current best evidence’, but also by often controversial judgements about ‘cost-effectiveness’.

² As is well known, ‘evidence-based medicine’ is itself a contested term, with many practitioners preferring the phrase ‘evidence-informed practice’.

Finally, as David Armstrong shows in this issue, EBM as a movement must also be placed within the context of medical professionalization, where the rise of EBM can be linked to a shift from a form of collegiate control of autonomy to one of state oversight, one in which clinical judgement is no longer controlled solely by the expert doctor, but also by algorithms of ‘safety’, ‘effectiveness’, ‘efficiency’ and ‘cost-effectiveness’. From this perspective, EBM can be understood as a means of curtailing not only costs, but also some of the many iatrogenic effects that have been attributed to a form of ‘eminence-based’ medicine, where doctors risk causing more harm than good.

Social studies of RCTs

It is through their scrutinizing of historical shifts, geographical variations, and political and commercial contexts of evidence-based medicine that social scientists, building on earlier work by clinicians themselves, have provided a new analytic perspective and voice. Some early milestone studies here include Harry Marks’ analysis of the role of twentieth-century therapeutic reformers in developing the RCT as a standardized technique for controlling external and internal experimental biases (1997), and Stefan Timmermans and Marc Berg’s study of the ways in which the RCT process needed to adjust to meet increasingly stringent regulatory requirements for determining drug efficacy and safety (Berg, 1995; Timmermans and Berg, 2003). Anthropologists, such as Vinh-Kim Nguyen (2005) and Adriana Petryna (2005), have ably explored the recent growth of trials conducted in developing regions, and the ways in which they are conducted according to a rather different set of ethical and regulatory demands than is typical in the West. Other social scientists have looked at economic aspects of RCTs, such as the exploitation of the medically vulnerable as research subjects (e.g. Trudy Dehue [2002], in her analysis of heroin users in RCTs in the Netherlands), gender issues (e.g. Oonagh Corrigan’s [2002] study of the privileging of male over female patients in phase II studies) and policy questions (e.g. Carl May’s [2006] ethnographic analysis of health technology assessment [HTA] exercises in the UK). In philosophy, Richard Ashcroft (2004) has explored the epistemological bases of RCTs, and, in bioethics, Kenneth Goodman (2003) has explored the ethical implications of the increasing use of RCT evidence among clinicians in everyday practice.

Through their focus on the politics and ethics of clinical trial construction, these last analyses point to one of the roles of social scientists and ethicists in building on work by practitioners themselves. And, although the development and use of RCTs has been extensively explored by numerous practitioners, scientists and historians of medicine (see, for example, Armitage, 1995; Chalmers *et al.*, 2002; Doll, 2003; Kaptchuk, 1998a, 1998b; Littlefield, 1982; Rees, 1997; Vandenbroucke, 2002), few of these analyses have investigated the political significance of RCTs, for example by exploring their rhetorical authority in mobilizing opinions about the efficacy and safety of drugs. The contributions in this volume have sought in part to pose such political questions, by examining the effects that the development of a global clinical trial infrastructure has had on health delivery, or by asking how questions such as ‘Does the drug work?’ or ‘Is it safe?’ are practically posed and investigated in clinical trials of drugs.

Andrew Lakoff's article, for example, seeks to understand the commercial and scientific motives for eliminating patients who could distort the results of antidepressant trials. Though Lakoff's analysis adds to the work of those who have illuminated the ways that RCTs are socially and politically constructed, his aim is not to be dismissive or merely critical, but rather to understand the various motivations of biotechnology firms in seeking, as his title indicates, 'the right patients for the drugs'. Lakoff's article implicitly adheres to the appeal offered earlier from Rose, and the suggestion that, rather than simply dismissing something as socially constructed, one should empirically explore the regulatory and political regimes through which a new technology or style of reasoning comes to be seen as the most morally sound and economically promising strategy.

This kind of empirical attention is needed in order to respond to the criticism of some proponents of EBM, that sociologists have been overly dismissive of EBM while failing adequately to grasp its diverse goals and motivations. On the other hand, even if there is some justice to the objection that some sociologists have been both too abstract and uniformly suspicious in their approach to EBM, the argument can equally be made that some proponents of evidence-based medicine have been too hasty in suggesting that the tools and methodologies of EBM are indispensable for conducting any manner of social or medical investigation. In an example, Jeanne Daly and colleagues (2006) propose a taxonomy of qualitative research, with single case studies at the lowest band of relevance and generalizable studies at the apex of the hierarchy. But is that the last word on the matter? Or could it be, as Nancy Cartwright argues in this issue, that it is more productive to keep in mind that 'what kind of method can do best at providing the information you want reliably, will differ from case to case'? Clearly, some kind of thoughtful middle ground is needed.

Gold standards

Are we so sure that randomized controlled trials are the 'best we've got' when it comes to evaluating whether or not a treatment does more good than harm? There are a number of components to this question of whether, and when the RCT is properly seen as the gold standard of EBM. On the one hand, there are debates concerning the 'taintedness' of the evidence-producing process, which often point to the distorting role of the pharmaceutical industry when lobbying for a relaxation of safety criteria, or when selectively publishing trial data in order to present a 'rosier' picture of the safety and efficacy of their drugs. For example, John Abraham argues in this issue that with pharmaceutical R&D costs soaring in a post-antibiotic world, corporate pressure has been piled on regulatory agencies such as the Food and Drug Administration (FDA) in America or the Medicines and Healthcare Products Regulatory Agency (MHRA) in the United Kingdom to speed up regulatory reviews of safety and efficacy data to the detriment of more robust safety requirements. In analysing recent trilateral (US, EU and Japan) efforts to harmonize safety standards internationally, Abraham (this volume) argues that this process has ultimately 'relaxed the extent to which trials might detect safety problems', in part by reducing the long-term exposure standard for assessing the safety of drugs intended for non-life-threatening disease from one year to six months. In this line of critique, the RCT, as Wendy Rogers (this volume) puts it in her review of Cochrane's *Effectiveness and efficiency*, 'has increasingly become a hand-maiden to the pharmaceutical industry'.

But these debates, Simon Wessely (this volume) points out, ‘do not represent a fundamental challenge to the status of the RCT, any more than detecting doctors who are less than competent or less than honest undermines the legitimacy of the medical enterprise’. The point being that allegations of fraud or undue influence do not question the ability of an RCT to produce reliable data, rather they point to weaknesses in, for example, regulatory oversight mechanisms, clinical trial protocols or outcome measures. For Wessely, the only claim to fame for the RCT is its ability to rule out confounders—even ones that we are unaware of—when evaluating the effectiveness of a therapy. There is simply no other way to be sure that it is the treatment in question and not some other factor (e.g. ‘regression to the mean’, ‘non-specific’ effects, placebo effect) that is ‘working’. Wessely argues: ‘It is not enough for a doctor to say that his or her treatment works, and nor is it enough for a patient to say likewise.’

Nancy Cartwright (this volume), on the other hand, argues that there is no gold standard when it comes to methods: ‘Gold methods are whatever methods will provide the information you need, reliably, from what you can do and from what you can know on the occasion.’ RCTs are examples of what Cartwright calls ‘clinchers’, methods for warranting causal claims that may well clinch the conclusion but at the cost of narrowness of scope. She argues that the formal methodology that clinical trial protocols must adhere to ‘puts severe constraints on the assumptions a target population must meet to justify exporting a conclusion from the test population to the target’. As Catherine Will points out in her article on ‘the alchemy of clinical trials’ (this volume), there is a constant negotiation between ‘rigour and relevance’ in the construction of randomized controlled trials, and disagreements between those constructing, carrying out and interpreting clinical trials are not necessarily always resolved in the process.

Taking another perspective, Uffe Juul Jensen (this volume) reminds us that, in controversies surrounding EBM and RCTs, ‘discussions are usually about the adequacy of different examples of evidence (such as controlled trials), while the meaning and function of the *concept of evidence* are rarely the focus of serious or systematic discussions’. By revisiting the work of Alvan Feinstein and Archie Cochrane as proponents of ‘patient-centred’ and ‘community-oriented’ medicine respectively, Jensen argues that evidence can be thought of in two distinct, yet inseparable senses. For Feinstein, diseases are things that are made visible or ‘demonstrated by new technology’, whereas illness is the clinically observed interaction between host and disease. For Cochrane, ‘states of health or illness are expressions of the success or failure experienced by the organism in its effort to respond adaptively to environmental challenges’. And where, in ‘patient-centred medicine’, evidence is gathered through clinical observations of the ways in which different patients interact in different ways with the same disease, in ‘community-centred medicine’ evidence is gathered by observing which treatments allow for an organism to respond most effectively to an external challenge.

Does it work?

Does the drug work? It is hard to imagine being confronted with a more succinct problem when it comes to a particular drug or therapy. Yet, as will become clear in the articles in

this issue, this is a problem for which clear-cut answers continue to remain elusive in many ways. As technologies, randomized controlled trials must rely on a number of different components if they are to render reliable and applicable results. First of all, a certain, standardized drug or therapy for which claims are made that it ‘alters the natural history of a disease for the better’ (Cochrane, 1972: 20) is required. It is this drug or therapy, as well as claims for its safety and efficacy, that are on trial in an RCT. Second, certain circumscribed and relatively homogeneous patient populations are required for the drug or therapy to be tested on—a component which has become increasingly difficult to satisfy in recent years. With the number of drugs being trialled at any given time on the increase, clinical trialling has become a global industry of its own in recent years, making patient recruitment a task that specialized clinical research organizations (CROs) are increasingly contracted by pharmaceutical companies to carry out on their behalf in accordance with ethical guidelines concerning informed consent. These research organizations also conduct the clinical trials and are responsible for adhering to specified trial protocols, which include diagnostic criteria for patient recruitment, instructions on double-blind placebo administration, rating scales and clinical outcome measures.

It is this process of clinical trialling that authors in this volume follow, illuminating the complexities inherent in this scientific ‘gold standard’. As is perhaps common in social scientific studies of scientific processes, the details of the clinical trial process are unpacked and scrutinized. We learn not to take anything for granted. Lakoff (this volume), in an analysis of how drug developers manage the uncertainties that can arise in depression trials, shows us how difficult securing consistent patient populations for clinical trials can be, especially when dealing with diseases where, as Cochrane put it, ‘improvement or deterioration has to be measured subjectively’ (1972: 24). Indeed, so difficult is it, Lakoff suggests, that ‘from the vantage of antidepressant development it is actually the drug, rather than the depressed patient, that serves as a stable reference point’. In this way, the primary problem for clinical trialists becomes not one of finding out whether a drug ‘works’ in a patient, but rather of figuring out how to design patient inclusion criteria and clinical outcome measures that will pick up a drug’s signal.

Such signals can be even more difficult to detect when dealing with so-called ‘complex interventions’. Catherine Will (this volume) shows how the construction of trials to test the effectiveness of lifestyle modification alongside preventive medication in the prevention of cardiovascular disease involves ‘considerable play of judgement around the choice of the object of randomization and of the measurements taken to stand for benefit or harm’ as clinicians and statisticians struggle to satisfy requirements of both rigour and relevance. Will argues that trial subject inclusion criteria are often negotiated in phase IV effectiveness trials ‘as a way to improve the likelihood of finding a strong relationship between the hoped-for benefit and the drug’.

Yet another problem facing clinical trialists today, as Adriana Petryna (this volume) shows, is that of ‘treatment saturation’ (‘People in the West live on pills’, as one of her informants put it). Petryna argues that such saturation ‘is making Americans and Western Europeans increasingly unusable from a drug-testing standpoint’, and as such she follows the clinical trial trail to middle- and low-income countries in an ethnography of the globalized clinical trial. Contract research organizations, which historically originated in America, are increasingly accessing pools of ‘treatment-naïve’ human subjects by offshoring clinical

trials to lower-income countries. This trend has in turn led to a market for clinical data which, problematically, has in some countries led to a situation where clinical trials (where participants receive ‘treatment’ free of charge) are in fact plugging health delivery gaps.

EBM and indeterminacy

As noted earlier, the life of a clinical trial does not end upon conclusion of the study. The data it generates has to be organized, interpreted and presented so as to become a part of an evidence base. In an article that moves away from the everyday practicalities of clinical trial construction and regulation, and towards the reception and deployment of evidence-based medicine and clinical trial results by practitioners, David Armstrong (this volume) provides us with a historical account of the rise of EBM as yet another strategy to address the indeterminacy and uncertainty that has always plagued doctor–patient encounters. With doctors facing numerous challenges to the autonomy of their medical profession from both patients and the state by the mid twentieth century, Armstrong suggests that EBM emerged as a new strategy for controlling indeterminacy which would appease both patients and health-care funders: ‘For patients, EBM promised the best available treatment; for governments, it offered the lure of elimination of wasteful—and often expensive—treatments.’ Armstrong goes on to conclude that ‘despite its rhetoric EBM does not dispense with indeterminacy’ having merely shifted it into another arena: the interpretation of clinical trial results.

In an interview with Linsey McGoey (this volume), problems of interpretation are taken up by Tim Kendall, joint-director of the National Collaborating Centre for Mental Health, a lead partner in developing mental health guidelines for NICE. Kendall discusses perhaps one of the most controversial issues surrounding the interpretation of clinical trial results, namely that of disclosure and publication bias. Kendall elaborates on some of the challenges he and his colleagues faced in collating and summarizing ‘current best evidence’ for a set of NICE clinical guidelines on the treatment of depression in a paediatric population, especially when trying to get access to unpublished clinical trials held by pharmaceutical companies, something that is a problem because ‘if you don’t have the full set of data there’s a risk that you would make the wrong decision, and you would make the wrong recommendation to people’.

At the same time, as Richard Ashcroft (this volume) points out in a review of Cochrane’s 35-year-old monograph—intended to both juxtapose and complement the review by Wendy Rogers—having access to the full set of data is not just important when making decisions about individual treatment but also for ‘getting rid of treatments known not to work, so as to free resources for treatments that are known to work, and for research to carry out further evaluations of treatments about which there is uncertainty’. Interpretation of clinical trial results has consequences not just for individuals, but also for health delivery systems.

An elusive evidence base

To conclude, what the collection of articles in this special issue clearly demonstrate is that answers to the questions of ‘Does a drug work?’ and ‘Is it safe?’ are highly complex and

rarely definitive. Instead they are open to much controversy and interpretation, which in turn feeds into debates about how randomized controlled trials should be constructed and regulated. Evidence bases can be elusive if there is insufficient disclosure or even deliberate suppression of available clinical trial data. The ideally ‘treatment-naïve’ populations that are so sought after by clinical trialists, have become quite rare in the Western world it seems, driving CROs to look East and South for larger pools of human subjects. Conductors of clinical trials struggle to get consistent patient populations that will respond reliably to a ‘drug’s signals’. Developers of rating scales grapple with the task of how best to measure and monitor ‘subjective’ improvement over time. The increasing relevance of ‘complex interventions’ has made trial design a constant negotiation between rigour and relevance.

Indeed, as many authors point out, in cases of the chronic, non-life-threatening diseases that pharmaceutical companies are increasingly targeting, the numbers needed to treat (NNT) have become so high that larger and larger clinical trials are required if a statistically significant result when compared to placebo is to be achieved. And finally, the ‘placebo effect’ itself has come to account for such a large proportion of treatment response that it has become more and more challenging to establish just what it is that is ‘working’.

It is perhaps telling that there are not many solutions provided by authors for the problems raised in this issue. It is difficult to conceive that anybody would be against attempts to identify the ‘best’, ‘most effective’ and ‘most safe’ forms of treatment for different diseases or conditions, especially when we know that health-care resources are finite. However, as the articles that follow so vividly capture, it is in coming to agreement over just what constitute the ‘best’ methods, regulations and protocols for identifying the safest and most effective treatments that controversies and debates have arisen. Should we be as pessimistic as John Abraham, who suggests that the commercialization of pharmaceutical development has led to an academic, industrial and regulatory complex that is ‘frequently inconsistent with furthering patients’ well-being and public health’? Or should we instead take Tim Kendall’s view when he argues that:

When we are deciding about which treatments work best, as compared to others—which is an important clinical question—then at this point in time, RCTs are probably the best way of answering that question. There are too many flaws in uncontrolled, non-randomized research. So, I am in favour of RCTs. No matter how problematic they are in the way they’re conducted, even when they’re done on a shoestring (which they often are if they’re not funded by a drug company), they are probably the best way that we currently have of deciding which treatments work and which ones don’t.

References

- Armitage, P. (1995). Before and after Bradford Hill: Some trends in medical statistics. *Journal of the Royal Statistical Society*, 158, 143–153.
- Armstrong, D. (2002). Clinical autonomy, individual and collective: The problem of changing doctors’ behaviour. *Social Science & Medicine*, 55, 1771–1777.
- Ashcroft, R. (2004). Current epistemological problems in evidence-based medicine. *Journal of Medical Ethics*, 30, 131–135.
- Berg, M. (1995). Turning a practice into a science: Reconceptualizing postwar medical practice. *Social Studies of Science*, 25, 437–476.

- Chalmers, I. (1990). Underreporting research is scientific misconduct. *Journal of the American Medical Association*, 263, 1405–1408.
- Chalmers, I., Hedges, L., & Cooper, H. (2002). A brief history of research synthesis. *Evaluation & the Health Professions*, 25, 12–37.
- Cochrane, A. (1972). *Effectiveness and efficiency: Random reflections of health services*. London: Nuffield Provincial Hospitals Trust.
- Cochrane, A. (1987). Cochrane in conversation with Max Blythe at Rhoose Farm, South Wales. Interview transcript, Medical Science Video Archive MSVA 024. Royal College of Physicians and Oxford Brookes University.
- Corrigan, O.P. (2002). ‘First in man’: The politics and ethics of women in clinical drug trials. *Feminist Review*, 72, 40–72.
- Daly, J., Willis, K., Small, R., Green, J., Welch, N., Kealy, M. *et al.* (2006). A hierarchy of evidence for assessing qualitative health research. *Journal of Clinical Epidemiology*, 60, 43–49.
- Dehue, T. (2002). A Dutch treat: Randomized controlled experimentation and the case of heroin-maintenance in the Netherlands. *History of the Human Sciences*, 15, 75–98.
- Doll, R. (2003). Fisher and Bradford Hill: Their personal impact. *International Journal of Epidemiology*, 32, 929–931.
- Goodman, K.W. (2003). *Ethics and evidence-based medicine*. Cambridge: Cambridge UP.
- Healy, D. (2001). The dilemma posed by new and fashionable treatments. *Advances in Psychiatric Treatment*, 7, 322–327.
- Jørgensen, A.W., Hilden, J., & Gøtzsche, P.C. (2006). Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: Systematic review. *British Medical Journal*, 333, 782.
- Kaptchuk, T. (1998a). Intentional ignorance: A history of blind assessment and placebo controls in medicine. *Bulletin of the History of Medicine*, 72, 389–433.
- Kaptchuk, T. (1998b). Powerful placebo: The dark side of the randomised control trial. *The Lancet*, 351, 1722–1725.
- Littlefield, A.M. (1982). *Ceteris paribus*: The evolution of the clinical trial. *Bulletin of the History of Medicine*, spring, 1–18.
- Marks, H. (1997) *The progress of experiment*. Cambridge: Cambridge University Press.
- May, C. (2006). Mobilising modern facts: Health technology assessment and the politics of evidence. *Sociology of Health and Illness*, 28, 513–532.
- Moncrieff, J., & Cohen, D. (2004). Rethinking models of psychotropic drug action. *Psychotherapy and Psychosomatics*, 970, 145–153.
- Mykhalovskiy, E., & Weir, L. (2004). The problem of evidence-based medicine: Directions for social science. *Social Science & Medicine*, 59, 1059–1069.
- Nguyen, V.-K. (2005). Antiretroviral globalism, biopolitics and therapeutic citizenship. In Ong, A., & Collier, S.J. (Eds), *Global assemblages: Technology, politics, and ethics as anthropological problems*. Oxford: Blackwell.
- Petryna, A. (2005). Drug development and the ethics of the globalized clinical trial. *Working Paper*, School of Social Science (New School for Social Research), Princeton, NJ, October.
- Rees, W.L. (1997). The place of controlled trials in the development of psychopharmacology. *History of Psychiatry*, 8, 1–20.
- Rose, N. (1999). *Governing the soul: The shaping of the private self*. London: Free Association Books.
- Sackett, D. (1996). Evidence based medicine: What it is and what it isn’t. *British Medical Journal*, 312, 71–72.
- Timmermans, S., & Berg, M. (2003). *The gold standard: The challenge of evidence-based medicine and standardization in health care*. Philadelphia, PA: Temple University Press.
- Vandenbroucke, J. (2002). The history of confounding. *History of Epidemiology*, 47, 216–224.