MEDICINE AND SOCIETY

Assessing the Gold Standard — Lessons from the History of RCTs

Laura E. Bothwell, Ph.D., Jeremy A. Greene, M.D., Ph.D., Scott H. Podolsky, M.D., and David S. Jones, M.D., Ph.D.

Over the past 70 years, randomized, controlled trials (RCTs) have reshaped medical knowledge and practice. Popularized by mid-20th-century clinical researchers and statisticians aiming to reduce bias and enhance the accuracy of clinical experimentation, RCTs have often functioned well in that role. Yet the past seven decades also bear witness to many limitations of this new "gold standard." The scientific and political history of RCTs offers lessons regarding the complexity of medicine and disease and the economic and political forces that shape the production and circulation of medical knowledge.

THE RISE OF RCTs

Physicians and medical researchers have attempted for millennia to evaluate therapeutic interventions with the use of case reports, case series, public demonstrations, testimonials, clinical reasoning, and occasionally clinical trials. As the role of science in medicine expanded in the late 19th century, physicians approached clinical research with increasing rigor. By the early 20th century, innovators had introduced many clinical-trial techniques to eliminate bias, including blinding, alternate assignment to trial groups, and statistical analysis.[1,2] When British epidemiologist Austin Bradford Hill formalized RCT methods in the 1940s, he built on many of these earlier strategies. Hill's work also propitiously coincided with Britain's investment in collaborative research during and after World War II. The Medical Research Council, for instance, provided a newly expanded infrastructure that could support RCTs.[3]

RCTs initially received mixed reviews. Some critics worried about the ethics of withholding promising new interventions from control groups.

Trialists countered that RCTs could determine whether new interventions were superior to the standard of care given to control groups.[4] Others argued that RCTs were urgently needed to assess manufacturers' claims about the flood of new medications — including antibiotics, antihypertensives, and antipsychotics — that emerged in the 1950s.[5,6] As an editorialist in the *Journal* cautioned in 1956, "Physicians should be particularly careful in accepting drugs purely on the basis of the manufacturer's evidence or on the basis of testimonials provided to the manufacturer. They should demand clear, unbiased, well studied and adequately controlled evidence produced and interpreted by reliable observers."[7] RCT proponents increasingly won over detractors. Soon, the U.S. National Institutes of Health and other government entities joined Britain in funding RCTs (Fig. 1).

Outside these academic and government circles, however, support for RCTs was initially weak. Pharmaceutical producers were reluctant to devote resources and time to RCTs when they could rely on expert testimonials and case reports to make broader claims about products.[3] The instability of this unregulated system became tragically apparent in 1961 when thalidomide, which had been given to thousands of pregnant women, was identified as the cause of an international epidemic of stillbirths and phocomelia. In response, the U.S. Congress enacted the Kefauver–Harris Amendments to the Food, Drug, and Cosmetic Act in 1962, mandating that new drugs be proven efficacious in "adequate and well-controlled investigations."[8] By 1970, the Food and Drug Administration (FDA) interpreted the amendments as requiring RCTs for the approval of new pharmaceuticals.[9]

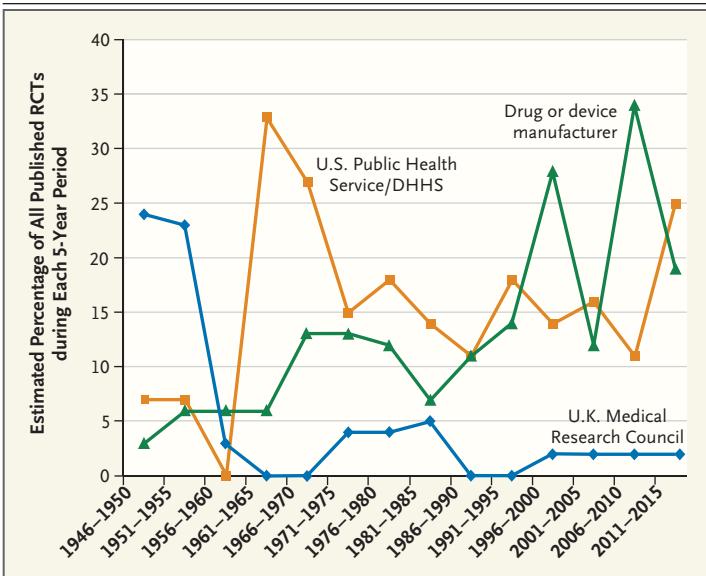These requirements, combined with postwar

Figure 1. Funding Sources for Randomized, Controlled Trials (RCTs).
Many early published RCTs were funded by government agencies in Britain and the United States, but the number of countries providing funding grew over time. Industry-funded RCTs expanded after regulators began requiring clinical trials for drug approval. As industry-sponsored trials proliferated, the number of published trials with undisclosed funding sources also increased. These trends reflect only the published literature; data are from a systematic sample of more than 600 published RCTs.[3] DHHS denotes Department of Health and Human Services.

growth of the U.S. pharmaceutical industry, contributed to the emergence of the United States as the leading producer of RCTs (Fig. 2).[3] The Council of the European Economic Community, the Japan government, and many national regulatory agencies soon implemented similar regulations. Over time, national regulators collaborated to establish international standards for clinical research, further systematizing RCTs.[10] In turn, to comply with regulations and obtain regulatory approval of new drug indications in a competitive marketplace, the pharmaceutical industry became a leading sponsor of RCTs. By the 1990s, industry had replaced governments and academic medicine as the primary producer of RCTs (Fig. 1).[3]

Clinical epidemiologists, meanwhile, promoted RCTs as the best means to make medicine more rational.[11,12] By the early 1980s, they had labeled RCTs the gold standard of medical knowledge.[13] As evidence-based medicine rose to prominence in ensuing decades, methodologic hierarchies emerged, with case reports at the bottom and RCTs at the top.

THE GOLD STANDARD THAT WASN'T

Yet RCTs have never monopolized medical knowledge production. A quick scan of the medical literature reveals that older methods, including case series and even case reports, continue to be valuable.[14-16] New methods of observational research continue to emerge — for instance, using large databases of patients to produce comparative effectiveness data on various treatment outcomes relatively efficiently in settings of routine care.[17,18] Physicians also continue to rely on physiological rationales in addition to empirical data. Coronary angioplasty and the stents that followed rose to prominence thanks not to a successful RCT but to the intuitive logic of the techniques and the compelling visual evidence provided by angiography.[19]

Even as RCTs have become standard in pharmaceutical research, clinical researchers have struggled to apply them to other areas of medicine. Although psychiatrists have conducted many RCTs of psychotherapy, critics have argued that it is inappropriate, and sometimes impossible, to evaluate such long-term, highly individualized interventions in that way.[20] Some major psychotherapy trials have been undermined by methodologic concerns.[21,22] Furthermore, because RCTs are more feasible for psychotropic drugs than for psychotherapy, the evidence base for psychotropics has become disproportionately more robust. Though that difference has benefited pharmaceutical manufacturers, it may also contribute to the use of less comprehensive approaches to psychiatric care.[3,23]

Surgical RCTs have faced similar complications. Surgeons began conducting RCTs in the 1950s — for instance, using sham controls to test the efficacy of internal-mammary-artery ligation for the treatment of angina pectoris.[24] As more surgical RCTs appeared in the 1960s and 1970s, however, surgeons increasingly recognized their limitations: each patient had unique pathological findings, each surgeon had different skills, and each operation involved countless choices about anesthesia, premedication, surgical approach, instrumentation, and postoperative care, all of which defied the standardization that clinical trials required.[25] Sham controls could not be used for major operations, which limited opportunities for blinded trials.

Such concerns played out in debates about

RCTs for coronary-artery bypass grafting (CABG). When the first major RCT of CABG revealed that most patients with chronic stable angina received no survival benefit from CABG,[26,27] critics pounced: the participants were too healthy, the surgeons too inexperienced, the operative mortality too high, and the statistical analysis suspect.[28,29] Prominent surgeons argued that RCTs were inappropriate for surgery.[30] René Favaloro, who had played a key role in developing CABG, argued that "randomized trials have developed such high scientific stature and acceptance that they are accorded an almost religious sanctification. . . . If relied on exclusively they may be dangerous."[31]

One long-standing, possibly intractable, concern has been the discrepancy between the time frame of RCTs and the fast pace of innovation. In debating how best to evaluate CABG in 1976, surgeons complained that "just when we have accumulated enough data over a sufficient time period, we find that surgical technique has improved or medical therapy changes, or both, and conclusions no longer apply."[32] Major RCTs have often required many years for patient enrollment, follow-up, and analysis. In cases of rapidly evolving therapies, RCT results have seemed outdated before they were published. When the COURAGE (Clinical Outcomes Utilizing Revascularization and Aggressive Drug Evaluation) trial showed disappointing efficacy results for coronary angioplasty in 2007,[33] the procedure's advocates argued that the results were no longer relevant because the bare-metal stents tested in the trial had been replaced by newer drug-eluting stents.[34] This logic, which assumes the superiority of any innovation, has created a setting in which trialists struggle to keep up with continuous innovations, similar to the "Red Queen" effect in evolutionary biology.[35]

Even well-conducted RCTs sometimes failed to influence medical practice. In the late 1960s, the meticulously designed University Group Diabetes Program trial linked the antidiabetic drug tolbutamide with increased cardiovascular mortality. Yet tolbutamide prescriptions paradoxically increased as controversies over the trial's conduct and interpretation persisted for more than a decade.[36,37] A similar scenario occurred when the publicly funded ALLHAT (Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial) revealed in 2002 that generic
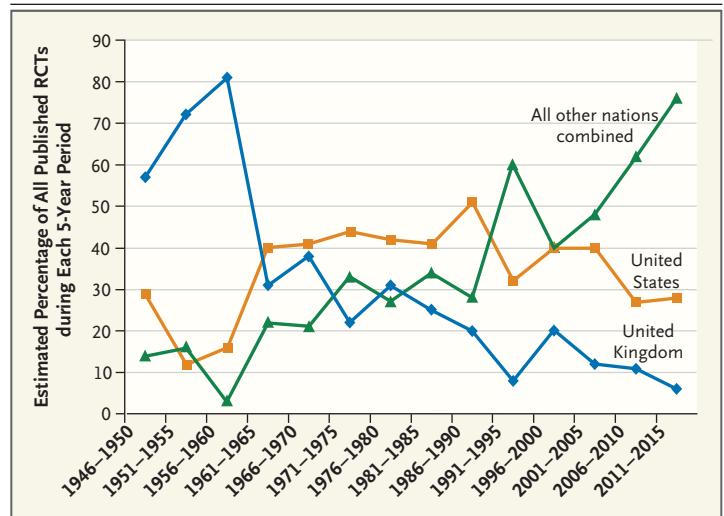

Figure 2. Locations of RCT Research Sites, 1946–2015.

Trends in the location of published RCTs reflect the British origins of the method and the U.S. enthusiasm for RCTs. After World War II, as the U.S. National Institutes of Health began supporting many trials, the U.S. pharmaceutical industry expanded, and the Kefauver–Harris amendments were passed, large numbers of RCTs were based in the United States. Many trials conducted outside the United States and the United Kingdom were spearheaded by U.S. or U.K. researchers and funders, particularly in the earlier years. As more international regulators began requiring RCTs for drug approval and RCTs increasingly became a global gold standard, trial sponsorship diversified. Data are from Bothwell.[3]

thiazide diuretics were as effective as newer, expensive calcium-channel blockers and angiotensin-converting–enzyme inhibitors in treating hypertension. As these findings were contested by pharmaceutical manufacturers and skeptical physicians, sales of the newer antihypertensives grew faster than those of diuretics.[38] Another 2002 RCT — a sham-surgery–controlled trial — defied conventional wisdom by showing no benefit of arthroscopic débridement for chronic osteoarthritis of the knee.[39] Many orthopedic surgeons dismissed the results and continued performing the procedure, even as the findings were confirmed repeatedly.[40]

On the other hand, some RCT results have been accepted as fact but have later proved lacking in external validity. RCTs have their challenges, from establishing appropriate inclusion criteria to standardizing interventions and determining the most relevant outcomes. These limitations have prompted researchers to pursue other methods, which have had their own limitations.

Social and ethical concerns have also chal-

lenged the legitimacy of some RCTs. The AIDS crisis brought many tensions into stark relief in the late 1980s. Patients, frustrated that RCTs would delay approvals of antiretroviral drugs, demanded access before trials had been completed.[41] Clinicians felt conflicted between their roles as physicians and as scientists.[42] Activists won support for more flexible approaches to clinical research, including the use of surrogate end points, conditional FDA approvals, and parallel tracks to provide access to drugs outside of trials. Critics worried that the loosened standards undermined scientific rigor and encouraged a risky deregulatory agenda championed by the drug industry.

Ethical disputes erupted in the 1990s over RCTs of treatments for HIV infection that were conducted in developing countries, especially regarding whether the low standard of care in some countries justified using placebo controls when they would be considered unethical in Europe or North America.[43,44] *Journal* editor Marcia Angell condemned "slavish adherence" to prescribed RCT practices when it caused a "retreat from ethical principles."[45]

Such controversies attracted attention from social scientists and policy scholars. As sociologist Steven Epstein noted, RCTs had become "crucial sites for the negotiation of credibility, risk, and trust." When they take place in fraught medical, social, and political contexts, RCTs, "rather than settling controversies, may instead reflect and propel them."[46] Historian Harry Marks argued that RCTs must be understood not merely as scientific techniques but also as social events: "even the simplest RCT is the product of a negotiated social order, replete with decisions — some contested, some not — and with unexamined assumptions."[36] Even though RCTs were developed to produce generalizable, universal biomedical knowledge, they have remained deeply entangled in local social conditions, economics, and politics.

ECONOMICS AND GEOGRAPHY OF KNOWLEDGE PRODUCTION

RCTs have also unintentionally limited the producers of medical knowledge. When case reports constituted valid evidence of therapeutic efficacy, a single physician, drawing on clinical experience, could write an article that might change clinical practice. RCTs, however, required collaborative research with substantial support. Over time, RCTs have become massive bureaucratic and corporate enterprises, demanding costly infrastructure for research design, patient care, record keeping, ethical review, and statistical analysis. By the 21st century, a single phase 3 RCT could cost $30 million or more.[47] As a result, trial sponsors often hail from North America, Western Europe, or East Asia, even when studies are conducted elsewhere. Consequently, RCTs disproportionately reflect the interests of industrialized regions.[48] The high costs of RCTs have had other unintended consequences: they have been invoked as a justification for high prescription-drug costs in markets lacking price controls.[49] Simultaneously, policymakers have recently proposed changes to regulatory law, such as the 21st Century Cures Act, that would curtail the role of RCTs in drug approval in the name of increased efficiency.

Furthermore, in part because of high trial costs, researchers and their funders have had substantial interests in achieving positive trial results. Considerable evidence suggests that industry-funded trials are more likely to produce favorable outcomes than publicly funded trials.[50] In addition, by the 1990s, it became clear that positive results tended to be published more often than negative results, to the detriment of medical knowledge. Regulators and journal editors responded to these problems with efforts to improve the transparency of RCTs, requiring the disclosure of financial conflicts of interest and the registration of all clinical trials so that negative trials wouldn't simply disappear.[51-53]

As RCTs developed into high-cost, high-value marketing tools, a clinical trials industry burgeoned. Having emerged in the late 1970s, contract research organizations (CROs) have become a $25 billion industry.[54] They have contributed to a shift in principal investigators in U.S. trials away from physician-scientists in academic teaching hospitals and toward nonacademic physicians working in the private sector on a contract basis.[55] CROs have also looked overseas for participants who have not previously received treatment, in middle-income countries where conditions are conducive to research. Countries now compete to convince the pharmaceutical industry and CROs that their regulatory, clinical, and public health profiles provide ideal trial condi-

tions, even when the products being tested are unlikely to be made available to local populations after trial completion.[48] But as research sites have diversified, research targets have not: much clinical research remains focused on drugs that may have a limited impact on public health but substantial marketing potential in high-income countries. Tuberculosis, malaria, and other scourges of low-income regions receive much less attention. The growing role of industry in global knowledge production has raised profound ethical and policy questions regarding the extent to which modern RCTs serve public health.

RCTS PAST, PRESENT, AND FUTURE

By the turn of the 21st century, RCTs had achieved the status of gold standard for therapeutic evidence — but one with well-documented limitations. Physicians continue to pursue alternative methods of knowledge production that are faster or less expensive than RCTs, or that claim to answer questions that RCTs cannot. Yet beyond medicine, RCTs are increasingly emulated, even idealized. Health policy researchers look for rare settings in which randomization can be implemented or where it has occurred inadvertently, as with the Oregon Medicaid experiment.[56] Development economists have placed RCTs at the center of a new experimental approach, proclaiming their potential "to revolutionize social policy during the 21st century, just as randomized trials revolutionized medicine during the 20th."[57] The extension of RCTs into other fields has drawn familiar critiques. Economist Angus Deaton, for example, argues that RCTs "cannot automatically trump other evidence, they do not occupy any special place in some hierarchy of evidence, nor does it make sense to refer to them as 'hard' while other methods are 'soft.'"[58]

Yet despite their limitations, RCTs have revolutionized medical research and improved the quality of health care by clarifying the benefits and drawbacks of countless interventions. Clinical investigators, supported by government funding and empowered by FDA regulations, have used RCTs to advance clinical research theory and practice. Critics have become increasingly adept at ferreting out flaws in RCTs, forcing trialists to be more vigilant in their designs. From a historical perspective, the RCT is not a single or stable technique, but an approach that has evolved as physicians have revised and refined clinical research.

The idea that RCTs would be the only authoritative arbiter to resolve medical disputes has given way to more pragmatic approaches. Experimentalists continue to seek new methods of knowledge production, from meta-analyses to controlled registry studies that can easily include large numbers of diverse patients. Observational methods are seen as complementary to RCTs, and new forms of surveillance can embed RCTs into the structure of data collection within electronic health records. RCTs are now just a part — though perhaps the most critical part — of a broad arsenal of investigative tools used to adjudicate efficacy and regulate the therapeutic marketplace. This status may continue to evolve with the recent turn (back) to personalized or precision medicine. As medicine focuses on the unique pathophysiology and coexisting conditions of individual patients, the applicability of the generalized data produced by RCTs will come under intensified scrutiny.

We find ourselves at a crucial point in the history of RCTs. Originally designed to reduce bias in research, RCTs have become sites of conflicting interests that merit careful scrutiny. Pharmaceutical and device manufacturers pursue data that will allow them to market products to new populations. Practicing physicians desire reliable evidence regarding which treatments will most benefit their patients.[59] RCTs serve both objectives, as historical entities representing at once scientific, political, and economic developments. Understanding this complex history enables us to evaluate RCTs more critically and effectively. Looking forward, given the role of RCTs in broader inequalities of global health research, how can we ensure that future trials address questions of genuine significance to medicine and global public health? Managing these historically contingent dimensions of RCTs will be a fundamental test of the roles and responsibilities of academic investigators, industry researchers, and government officials who work to advance reliable and useful medical research.

2179

Harvard Medical School (L.E.B.), and the Department of Global Health and Social Medicine, Harvard Medical School (S.H.P., D.S.J.), Boston, and the Department of the History of Science, Harvard University, Cambridge (D.S.J.) — both in Massachusetts; and the Division of General Internal Medicine and the Department of the History of Medicine, Johns Hopkins University School of Medicine, Baltimore (J.A.G.).

1. Opinel A, Tröhler U, Gluud C, et al. Commentary: the evolution of methods to assess the effects of treatments, illustrated by the development of treatments for diphtheria, 1825-1918. Int J Epidemiol 2013;42:662-76.
2. Podolsky SH, Jones DS, Kaptchuk TJ. From trials to trials: blinding, medicine, and honest adjudication. In: Robertson CT, Kesselheim AS, eds. Blinding as a solution to bias: strengthening biomedical science, forensic science, and law. Amsterdam: Elsevier, 2016:45-58.
3. Bothwell L. The emergence of the randomized controlled trial: origins to 1980 (Ph.D. dissertation). New York: Columbia University Academic Commons, 2014.
4. Hill AB. Medical ethics and controlled trials. Br Med J 1963; 1:1043-9.
5. Podolsky SH. The antibiotic era: reform, resistance, and the pursuit of a rational therapeutics. Baltimore: Johns Hopkins University Press, 2015.
6. Goodman J. Pharmaceutical industry. In: Cooter R, Pickstone JV, eds. Companion to medicine in the twentieth century. New York: Taylor and Francis, 2000:150.
7. Antibiotic combinations. N Engl J Med 1956;255:1057-9.
8. Greene JA, Podolsky SH. Reform, regulation, and pharmaceuticals — the Kefauver–Harris Amendments at 50. N Engl J Med 2012;367:1481-3.
9. Carpenter D. Reputation and power: organizational image and pharmaceutical regulation at the FDA. Princeton, NJ: Princeton University Press, 2010.
10. Kuo W-H. Japan and Taiwan in the wake of bio-globalization: drugs, race and standards (Ph.D. dissertation). Cambridge: Massachusetts Institute of Technology, 2005.
11. Daly J. Evidence-based medicine and the search for a science of clinical care. Berkeley: University of California Press, 2005.
12. Timmermans S, Berg M. The gold standard: the challenge of evidence-based medicine and standardization in health care. Philadelphia: Temple University Press, 2003.
13. Jones DS, Podolsky SH. The history and fate of the gold standard. Lancet 2015;385:1502-3.
14. Baize S, Pannetier D, Oestereich L, et al. Emergence of Zaire Ebola virus disease in Guinea. N Engl J Med 2014;371:1418-25.
15. Varkey JB, Shantha JG, Crozier I, et al. Persistence of Ebola virus in ocular fluid during convalescence. N Engl J Med 2015; 372:2423-7.
16. Brasil P, Pereira JP Jr, Raja Gabaglia C, et al. Zika virus infection in pregnant women in Rio de Janeiro — preliminary report. N Engl J Med. DOI: 10.1056/NEJMoa1602412.
17. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol 2005;58:323-37.
18. Schneeweiss S. Developments in post-marketing comparative effectiveness research. Clin Pharmacol Ther 2007;82:143-56.
19. Gruntzig A. Transluminal dilatation of coronary-artery stenosis. Lancet 1978;1:263.
20. Seligman MEP. The effectiveness of psychotherapy: the Consumer Reports study. Am Psychol 1995;50:965-74.
21. Elkin I, Shea MT, Watkins JT, et al. National Institute of Mental Health Treatment of Depression Collaborative Research Program: general effectiveness of treatments. Arch Gen Psychiatry 1989;46:971-82.
22. Ablon JS, Jones EE. Validity of controlled clinical trials of psychotherapy: findings from the NIMH Treatment of Depression Collaborative Research Program. Am J Psychiatry 2002;159: 775-83.
23. Deacon BJ. The biomedical model of mental disorder: a critical analysis of its validity, utility, and effects on psychotherapy research. Clin Psychol Rev 2013;33:846-61.
24. Cobb LA, Thomas GI, Dillard DH, Merendino KA, Bruce RA. An evaluation of internal-mammary-artery ligation by a double-blind technic. N Engl J Med 1959;260:1115-8.
25. Love JW. Drugs and operations: some important differences. JAMA 1975;232:37-8.
26. Murphy ML, Hultgren HN, Detre K, Thomsen J, Takaro T, Participants of the Veterans Administration Cooperative Study. Treatment of chronic stable angina — a preliminary report of survival data of the randomized Veterans Administration cooperative study. N Engl J Med 1977;297:621-7.
27. Braunwald E. Coronary-artery surgery at the crossroads. N Engl J Med 1977;297:661-3.
28. Jones DS. Visions of a cure: visualization, clinical trials, and controversies in cardiac therapeutics, 1968-1998. Isis 2000;91: 504-41.
29. Special correspondence: a debate on coronary bypass. N Engl J Med 1977;297:1464-70.
30. Bonchek LI. Are randomized trials appropriate for evaluating new operations? N Engl J Med 1979;301:44-5.
31. Favaloro RG. Critical analysis of coronary artery bypass graft surgery: a 30-year journey. J Am Coll Cardiol 1998;31:Suppl:1B-63B.
32. Ullyot DJ, Wisneski J, Sullivan RW, Gertz EW. Reply to: The controversy over coronary arterial surgery: inappropriate controls, inappropriate publicity. J Thorac Cardiovasc Surg 1976;72: 945.
33. Boden WE, O'Rourke RA, Teo KK, et al. Optimal medical therapy with or without PCI for stable coronary disease. N Engl J Med 2007;356:1503-16.
34. Kereiakes DJ, Teirstein PS, Sarembock IJ, et al. The truth and consequences of the COURAGE trial. J Am Coll Cardiol 2007;50: 1598-603.
35. Jones DS. Therapeutic evolution or revolution? Metaphors and their consequences. In: Greene JA, Condrau F, Watkins ES, eds. Therapeutic revolutions: pharmaceuticals and social change in the twentieth century. Chicago: University of Chicago Press, 2016.
36. Marks HM. The progress of experiment: science and therapeutic reform in the United States, 1900-1990. Cambridge, United Kingdom: Cambridge University Press, 1997.
37. Greene JA. Prescribing by numbers: drugs and the definition of disease. Baltimore: Johns Hopkins University Press, 2007.
38. Pollack A. The minimal impact of a big hypertension study. New York Times. November 27, 2008:B1.
39. Moseley JB, O'Malley K, Petersen NJ, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. N Engl J Med 2002;347:81-8.
40. Thorlund JB, Juhl CB, Roos EM, Lohmander LS. Arthroscopic surgery for degenerative knee: systematic review and meta-analysis of benefits and harms. Br J Sports Med 2015;49:1229-35.
41. Epstein S. Impure science: AIDS, activism, and the politics of knowledge. Berkeley: University of California Press, 1996.
42. Hellman S, Hellman DS. Of mice but not men — problems of the randomized clinical trial. N Engl J Med 1991;324:1585-9.
43. Lurie P, Wolfe SM. Unethical trials of interventions to reduce perinatal transmission of the human immunodeficiency virus in developing countries. N Engl J Med 1997;337:853-6.
44. Rothman DJ. The shame of medical research. New York Rev Books 2000;47(19):60-4.
45. Angell M. The ethics of clinical research in the Third World. N Engl J Med 1997;337:847-9.
46. Epstein S. Activism, drug regulation, and the politics of

therapeutic evaluation in the AIDS era: a case study of ddC and the 'surrogate markers' debate. Soc Stud Sci 1997;27:691-726.

47. Sertkaya A, Birkenbach A, Berlind A, Eyraud J. Examination of clinical trial costs and barriers for drug development: report to the Assistant Secretary of Planning and Evaluation (ASPE). Washington, DC: Department of Health and Human Services, July 25, 2014.

48. Petryna A. When experiments travel: clinical trials and the global search for human subjects. Princeton, NJ: Princeton University Press, 2009.

49. Scannell J. Four reasons drugs are expensive, of which two are false. Forbes.com. October 13, 2015 (http://www.forbes.com/sites/matthewherper/2015/10/13/four-reasons-drugs-are-expensive-of-which-two-are-false).

50. Bourgeois FT, Murthy S, Mandl KD. Outcome reporting among drug trials registered in ClinicalTrials.gov. Ann Intern Med 2010;153:158-66.

51. Burdett S, Stewart LA, Tierney JF. Publication bias and meta-analyses: a practical example. Int J Technol Assess Health Care 2003;19:129-34.

52. Zarin DA, Ide NC, Tse T, Harlan WR, West JC, Lindberg DA. Issues in the registration of clinical trials. JAMA 2007;297:2112-20.

53. Boyd EA, Cho MK, Bero LA. Financial conflict-of-interest policies in clinical research: issues for clinical investigators. Acad Med 2003;78:769-74.

54. Bodenheimer T. Uneasy alliance — clinical investigators and the pharmaceutical industry. N Engl J Med 2000;342:1539-44.

55. Fisher JA, Kalbaugh CA. United States private-sector physicians and pharmaceutical contract research: a qualitative study. PLoS Med 2012;9(7):e1001271.

56. Baicker K, Taubman SL, Allen HL, et al. The Oregon experiment — effects of Medicaid on clinical outcomes. N Engl J Med 2013;368:1713-22.

57. The World Bank is finally embracing science. Lancet 2004;364:731-2.

58. Deaton A. Instruments, randomization, and learning about development. J Econ Lit 2010;48:424-55.

59. Avorn J. Powerful medicines: the benefits, risks, and costs of prescription drugs. New York: Alfred A. Knopf, 2004.

Copyright © 2016 Massachusetts Medical Society.