REVIEW ARTICLE

THE CHANGING FACE OF CLINICAL TRIALS

Jeffrey M. Drazen, M.D., David P. Harrington, Ph.D., John J.V. McMurray, M.D., James H. Ware, Ph.D., and Janet Woodcock, M.D., *Editors*

Evidence for Health Decision Making — Beyond Randomized, Controlled Trials

Thomas R. Frieden, M.D., M.P.H.

A core principle of good public health practice is to base all policy decisions on the highest-quality scientific data, openly and objectively derived.[1] Determining whether data meet these conditions is difficult; uncertainty can lead to inaction by clinicians and public health decision makers. Although randomized, controlled trials (RCTs) have long been presumed to be the ideal source for data on the effects of treatment, other methods of obtaining evidence for decisive action are receiving increased interest, prompting new approaches to leverage the strengths and overcome the limitations of different data sources.[2-8] In this article, I describe the use of RCTs and alternative (and sometimes superior) data sources from the vantage point of public health, illustrate key limitations of RCTs, and suggest ways to improve the use of multiple data sources for health decision making.

In large, well-designed trials, randomization evenly distributes known and unknown factors among control and intervention groups, reducing the potential for confounding. Despite their strengths, RCTs have substantial limitations. Although they can have strong internal validity, RCTs sometimes lack external validity; generalizations of findings outside the study population may be invalid.[2,4,6] RCTs usually do not have sufficient study periods or population sizes to assess duration of treatment effect (e.g., waning immunity of vaccines) or to identify rare but serious adverse effects of treatment, which often become evident during postmarketing surveillance and long-term follow-up but could not be practically assessed in an RCT. The increasingly high costs and time constraints of RCTs can also lead to reliance on surrogate markers that may not correlate well with the outcome of interest. Selection of high-risk groups increases the likelihood of having adequate numbers of end points, but these groups may not be relevant to the broader target populations. These limitations and the fact that RCTs often take years to plan, implement, and analyze reduce the ability of RCTs to keep pace with clinical innovations; new products and standards of care are often developed before earlier models complete evaluation. These limitations also affect the use of RCTs for urgent health issues, such as infectious disease outbreaks, for which public health decisions must be made quickly on the basis of limited and often imperfect available data. RCTs are also limited in their ability to assess the individualized effect of treatment, as can result from differences in surgical techniques, and are generally impractical for rare diseases.

Many other data sources can provide valid evidence for clinical and public health action. Observational studies, including assessments of results from the

From Atlanta, GA. The author is the former director of the Centers for Disease Control and Prevention. Address reprint requests to Dr. Frieden at tfrieden@gmail.com.

N Engl J Med 2017;377:465-75.
DOI: 10.1056/NEJMra1614394

implementation of new programs and policies, remain the foremost source, but other examples include analysis of aggregate clinical or epidemiologic data. In the late 1980s, the high rate of the sudden infant death syndrome (SIDS) in New Zealand led to a case–control study comparing information on 128 infants who died from SIDS and 503 control infants.[9] The results identified several risk factors for SIDS, including prone sleeping position, and led to the implementation of a program to educate parents to avoid putting their infants to sleep on their stomachs — well before back-sleeping was definitively known to reduce the incidence of SIDS. The substantial reduction in the incidence of SIDS that resulted from this program became strong evidence of efficacy; implementation of an RCT for SIDS would have presented ethical and logistic difficulties. Similarly, the evidence base for tobacco-control interventions has depended heavily on analysis of the results of policies, such as taxes, smoke-free laws, and advertising campaigns that have generated robust evidence of effectiveness — that is, practice-based evidence.

Current evidence-grading systems are biased toward RCTs, which may lead to inadequate consideration of non-RCT data.[10] Objections to observational studies include the potential for bias from unrecognized factors along with the belief that these studies overestimate treatment effects.[11] Although overestimation bias has been shown in some observational studies (e.g., overestimation of the effect of influenza vaccination on reducing mortality among older persons as a result of bias from healthy vaccine recipients[12]), comparisons of validity between observational studies and RCTs have dispelled many misperceptions.[4,6,13,14] A widely cited example involves the cardiovascular health risks associated with the use of menopausal hormone therapy. Data from an observational study suggested that menopausal hormone therapy would reduce the risk of heart disease[15]; results from a subsequent RCT showed increased cardiovascular risks.[16] Although initially these differences were thought to indicate weaknesses in the observational study, further analyses determined that both studies had valid results for their patient populations and that discrepancies were probably due to the timing of initiation of hormone therapy in relation to the onset of menopause.[17-21] If so, then the RCT and observational study showed similar findings. However, a broad recommendation to use hormone therapy was made prematurely. Determining when data are sufficient for action is difficult, but the bar should be much higher when recommending that millions of persons with no disease take medications. This line of reasoning does not suggest that the Food and Drug Administration should be less stringent in their review of drug safety and efficacy, but rather that there should be rigorous review of all potentially valid data sources.

No study design is flawless, and conflicting findings can emerge from all types of studies. The following examples show the importance of recognizing the strengths and limitations in all data sources and finding ways to obtain the most useful data for health decision making.

Validity of Alternative Data Sources — The Live Attenuated Influenza Vaccine

Rigorous analyses after the implementation of a public health program can provide critically important information, such as data on vaccine effectiveness. Analyses of influenza vaccination efforts are a prime example, because, unlike other vaccines, influenza vaccines are given and evaluated for effectiveness yearly. The ability of an influenza vaccine to prevent influenza-related illness is affected by many factors, including genetic changes in the virus as well as host factors including age, underlying medical conditions, and previous infections and vaccinations. In the United States, the effectiveness of the influenza vaccine is monitored through the Influenza Vaccine Effectiveness Network. These data are used to derive estimates of the number of influenza-related illnesses, hospitalizations, and deaths prevented each year through vaccination, which, in turn, provide critical information to help measure, evaluate, and guide public health interventions.

First licensed in 2003, the live attenuated influenza vaccine, known as the "nasal spray" influenza vaccine, has been approved for use in healthy children and adults 2 to 49 years of age since 2007.[22] The vaccine showed good protection for both adults and children in postlicen-

sure RCTs, and, in June 2014, on the basis of results from several RCTs showing superior efficacy of the live attenuated vaccine over the inactivated influenza vaccine in children,[23-25] the Advisory Committee for Immunization Practices (ACIP) issued a preference for its use in healthy children 2 to 8 years of age for the 2014–2015 influenza season.[26] A subsequent observational study of the effectiveness of the live attenuated and inactivated influenza vaccines, however, showed worse performance for live attenuated vaccine than was shown in the RCTs,[27] and the ACIP did not renew its preference for the live attenuated vaccine over inactivated vaccine in healthy children for the 2015–2016 season. More recently, on the basis of an observed vaccine efficacy for the live attenuated vaccine that was at or near zero, especially against the 2009 H1N1 pandemic influenza virus,[27-29] the ACIP recommended that the nasal spray vaccine not be used during the 2016–2017 influenza season.[30] In this example, changes in vaccine formulation (from trivalent to quadrivalent), the population vaccinated (e.g., natural immunity resulting in neutralization of live vaccine), or another factor or factors caused the RCT data to lack external validity and be misleading, as compared with prospectively collected vaccine-efficacy data. Future studies may provide clarification regarding the reasons for these differences, but both RCTs and observational data may be needed.

RELEVANCE TO PROGRAM CONDITIONS — DIRECTLY OBSERVED TREATMENT FOR TUBERCULOSIS

Although the use of a single drug in the 1946 RCT of streptomycin for the treatment of tuberculosis[31] rapidly led to resistance, the success of the trial spurred a series of long-term RCTs for tuberculosis treatment conducted over four decades by the British Medical Research Council with collaborators throughout the world.[32,33] Each trial built on previous findings, with the effect of refining drug regimens and minimizing the duration of antituberculosis treatment. The importance of directly observed treatment was realized as treatment moved from sanatoriums to homes.[34,35] The approach, implemented from 1958 forward,[33] evolved to directly observed treatment, short-course (DOTS), with standard, first-line regimens, and, for persons infected with multidrug-resistant strains, "DOTS-plus," involving second-line, reserve drugs.[36]

Studies have purported to show that directly observed treatment offers no advantage over self-administered treatment.[37,38] A limitation of these studies has been lack of evaluation of the health, epidemiologic, and societal costs of relapse or of the rare but devastating progression to drug-resistant tuberculosis. Although these studies have been conducted with intensive oversight, they have not established a method of treatment that can be consistently applied to a large program in which thousands or millions of patients are treated. In addition, an RCT for a tuberculosis treatment method would be unable to predict or account for the harms from the rare but catastrophic secondary, population-wide effects of development and spread of multidrug resistance.

Examples of non-RCT efforts to evaluate the effect of DOTS and DOTS-plus on multidrug-resistant tuberculosis include decision analyses of program effect,[39] genotyping of isolates from patients in communities with different directly observed treatment practices,[40] and reviews of medical and public health records along with epidemiologic and laboratory analyses of multidrug-resistant tuberculosis outbreaks.[41] These non-RCT studies have contributed to continued refinements in treatment and follow-up and reduced risks of resistance. For these and other reasons, the American Thoracic Society, World Health Organization, and Centers for Disease Control and Prevention continue to recommend directly observed treatment as the standard of practice.

POPULATION-WIDE ANALYSIS — THE EFFECT OF SODIUM INTAKE ON CARDIOVASCULAR HEALTH

Cardiovascular disease remains the leading cause of death in the United States.[42] A major risk factor for cardiovascular disease is hypertension, which currently affects approximately 29% of U.S. adults.[43] An important strategy for lowering blood pressure is reducing excess sodium intake, particularly through changes to the food supply.[44] A robust body of evidence, including an analysis

of more than 100 randomized trials, shows that reducing sodium intake reduces blood pressure among adults.[45] There is also evidence, based on trends at the population level, that reducing sodium intake prevents cardiovascular disease.[46] Meta-analysis of sodium-reduction trials of at least 6 months' duration in which moderate reductions in intake were achieved, as well as well-designed, long-term cohort studies, have provided strong evidence that lower sodium intake is associated with a reduced incidence of cardiovascular events.[47,48]

The benefits of sodium reduction have been questioned by some researchers on the basis of several studies that report a J-shaped relationship between sodium intake and cardiovascular outcomes.[49-51] These studies, however, have been shown to have methodologic flaws, including those related to the assessment of usual sodium intake, the potential for reverse causality, inadequate follow-up, residual confounding, and insufficient power.[52] Accurate assessment of long-term, usual sodium intake is critical in cohort studies that relate individual sodium intake to long-term outcomes and requires multiple 24-hour urine collections over a period of time.[52-54] Spot or single 24-hour urine collections have a high degree of intraindividual variation that may not be overcome by correction or large sample size.[54-56] Because of challenges in accurately measuring usual sodium intake and excretion and the potential for misclassification of exposure, cohort studies must use multiple 24-hour urine collections[48] to be valid, and study designs that use population means, which are subject to less variation than measurements of individual intake, often provide more reliable information.[57] This may be why studies that assess sodium intake and cardiovascular events on a population level have shown beneficial effects of sodium-intake reduction,[46] whereas studies with less accurate measures of individual intake have not.[54,57]

Even for established risk factors, RCTs can yield answers that are simply wrong. A well-known example is the large Multiple Risk Factor Intervention Trial (MRFIT) on cardiovascular disease, which showed insufficient differences in health outcomes resulting from interventions such as smoking cessation and exercise.[58] Although longer follow-up showed that the trial may have accurately identified benefits from smoking cessation and improvements in nutrition, the study highlighted problems in implementing and measuring the effects of substantial lifestyle changes — in particular, insufficient follow-up duration, possible adoption of interventions by participants in the comparison group, and inadequate adherence to recommended interventions by participants in the study population.

Although some researchers have called for large, long-term RCTs examining the effects of sodium-intake reduction on clinical outcomes to inform population-wide sodium-reduction efforts, this approach is similarly not feasible. Such trials would require tens of thousands of participants undergoing randomization to a high-sodium or low-sodium diet, with adherence to the intervention and follow-up of at least 5 years.[47] This study design is impractical, particularly given the challenges with adherence to a low-sodium diet in our current food environment. As with many other topics in public health, conflicting findings from studies that use different methods are to be expected. Critical analysis of study methods and measurement and examination of the totality of the evidence are essential in order to interpret results correctly and make appropriate recommendations for action.[59]

RARE DISEASES — THE IMPORTANCE OF DISEASE REGISTRIES AND OTHER METHODS

Approximately 5000 to 7000 conditions fit the definition of a rare disease, with more than 50 million people affected throughout the world.[60,61] Because of small sample sizes and logistic constraints, it is unlikely that RCTs will be performed for most of these conditions; actionable information may be most likely to be obtained from meticulous analysis of the treatment of different patients by different methods. Such an approach was used to determine that isoniazid, injectable medications, and fluoroquinolone antibiotic agents were most likely to lead to successful treatment for common strains of multidrug-resistant tuberculosis.[41] Despite the Orphan Drug Act, which was passed in 1983 to provide industry incentives for the development of clinical treatments for rare diseases, the op-

tions for most patients are limited. A movement to create a global rare-disease patient registry along with a centralized database of biorepositories for rare biospecimens followed from a 2010 workshop, sponsored by the National Institutes of Health, involving researchers, advocacy groups, and stakeholders.[62] The Rare Diseases Human Biospecimens/Biorepositories (RD-HuB) makes rare-disease specimens available to researchers and informs patients of ongoing studies. Although such registries could potentially lead to RCTs, attaining sufficient study-population sizes could remain an impediment. Alternately, these registries could be used to collect detailed case studies, including standardized information on individual treatment and clinical status, which could be used to enhance understanding of a particular disease and its treatment and improve the health of affected patients. For example, standardizing and aggregating data on clinical features, treatment, and outcomes from case reports and case series may reveal ways to improve diagnosis and treatment.

COSTS AND INFRASTRUCTURE — RELIABLE RESULTS FROM MORE FEASIBLE STUDY DESIGNS

Large observational studies, with longer follow-up, can be tailored to minimize bias in a manner analogous to the way bias is minimized in RCTs. In one such study, data from the Veterans Health Administration (VA) and Medicare were used to examine outcomes of treatment with sulfonylureas and thiazolidinediones — two second-line drugs for type 2 diabetes.[63] The study used physician-prescribing patterns to approximate an RCT: determinations were made for patients to receive a sulfonylurea or thiazolidinedione on the basis of how often their physician had prescribed the drugs during the previous year (i.e., patients of physicians who usually prescribed sulfonylureas were assigned to receive a sulfonylurea, and those whose physicians usually prescribed thiazolidinediones were assigned to receive a thiazolidinedione). With more than 80,000 patients monitored for up to 10 years, the study was 20 times larger and had a much longer follow-up than previous RCTs comparing the effectiveness of second-line diabetes drugs. The results showed a 68% higher risk of avoidable hospitalization and a 50% higher risk of death associated with treatment with sulfonylureas, as compared with thiazolidinediones, providing strong evidence-based information for clinical decision making while also avoiding many of the limitations of RCTs.

The VA is also undertaking a new type of randomized trial to compare the use of chlorthalidone versus hydrochlorothiazide for the treatment of hypertension.[64] Both medications, which are diuretics, have been used for more than 50 years, but more than 95% of the million or more veterans who are prescribed this type of diuretic receive hydrochlorothiazide, as compared with the 2.5% receiving chlorthalidone.[65] However, there is evidence that chlorthalidone, the older of the two drugs, is more effective in preventing cardiovascular events[66] and reducing mortality.[67,68] Using data from electronic medical records, with reliance on the patients' primary care physician instead of additional study personnel, the trial plans to enroll approximately 13,500 veterans older than 65 years of age who are currently receiving hydrochlorothiazide. These patients will then be randomly assigned to receive hydrochlorothiazide or chlorthalidone over a 3-year study period. This study design simplifies the infrastructure and greatly reduces the costs involved in a traditional, large RCT.[64] With approximately 50 million prescriptions for hydrochlorothiazide filled each year in the United States, even small reductions in cardiovascular events associated with chlorthalidone use that may be identified through this study would have a substantial effect in the prevention of cardiovascular disease.

MOVING FORWARD — OVERCOMING THE "DARK MATTER" OF CLINICAL MEDICINE

For much, and perhaps most, of modern medical practice, RCT-based data are lacking and no RCT is being planned or is likely to be completed to provide evidence for action. This "dark matter" of clinical medicine leaves practitioners with large information gaps for most conditions and increases reliance on past practices and clinical lore.[4,69,70] Elevating RCTs at the expense of other

Table 1. Selected Strengths and Weaknesses of Various Study Designs, along with Examples of Studies with Effects on Policy or Practice.

Data Source	Strengths	Weaknesses	Examples of Effects on Policy or Practice
Randomized, controlled trials	Can identify causal relationships Can reduce bias and confounding Can determine efficacy: can establish definitively which treatment methods are superior	Potential for validity to be limited to study population, with limited relevance to actual conditions Potential for surrogate markers, if used, to not correlate with outcome of interest Resource-intensive with regard to costs: high costs may lead to designs with inadequate sample size Resource-intensive with regard to time: completion may not occur until after introduction of new products or treatment methods, so that trials are not studying what is used in actual clinical practice Impractical for urgent situations and certain conditions (e.g., rare diseases) May not account for effects beyond study population (i.e., effects on persons not participating in the trial, such as spread of infection to others)	Trials have defined the cardiovascular benefits of lowering low-density lipoprotein cholesterol and of lowering blood pressure in various patient populations[75-77] Trials have established and continue to refine tuberculosis treatment regimens used globally[32,33]
Meta-analyses, systematic reviews, decision analyses	Can broaden capacity to test hypotheses and detect patterns and effects Allow for pooled results that can potentially yield more robust estimates Can adjust for underlying study rigor and sample size Do not require additional new data collection	Potential for invalid conclusions from the combination of different data sources; validity limited by quality of underlying studies and different methods of measuring the same outcome among studies; potential for false sense of precision Because many hypotheses can easily be tested, potential exists for introduction of systematic bias through selective publication of positive findings Limited availability of valid studies to analyze for some topics	Studies have evaluated factors associated with stroke, myocardial infarction, and death among patients with carotid artery stents[78] Studies have assessed effects of cholesterol-lowering medications and patient selection for this treatment[79,80] Studies have analyzed different approaches for prevention of cancer (e.g., among women with *BRCA1* or *BRCA2* mutations), for prevention of colon cancer, and for treatment of prostate cancer[81-83]
Prospective cohort studies	Establish temporal relationship Can evaluate a range of outcomes Can evaluate rare exposures Allow for nested studies	Inefficient for studying rare diseases Resource-intensive, since most cohorts must be followed for many years Potential for nonrepresentative study populations (e.g., persons who are less mobile) resulting from losses to follow-up	Identification of risk factors for breast and colon cancer, cardiovascular disease, hip fracture, eye disease, and decreased cognitive function (Nurses' Health Study) led to changes in screening, prevention, and treatment[84] Identification of risk of cancer and death among patients infected with hepatitis C virus led to intensified efforts to establish and provide effective treatment[85]
Retrospective cohort studies	Establish temporal relationship Can evaluate a range of outcomes associated with a given exposure Can evaluate rare exposures Allow for nested studies Can be conducted rapidly	Inefficient for studying rare diseases Resource-intensive (with regard to costs and time) Potential for difficulties in correcting for recall and other forms of bias and for confounding	Assessment of prognosis and treatment in different types of cancer led to better treatment protocols[86] Assessment of survivors of childhood cancer led to recognition of increased risk of post-treatment cardiac complications, enabling better clinical care[87]
Case–control studies	Efficient for studying rare outcomes and potential associated exposures Can be conducted rapidly and generally at low cost Can rapidly yield information with implications for action	Potential for varying quality of exposure assessment data May be more prone to bias than cohort studies because of selection bias and other study effects No information about rates of disease or temporal trends	Identification of risk factors for the sudden infant death syndrome (SIDS) led to intervention programs that have greatly reduced infant mortality[9] Determination of common exposures has led to identification of sources of infection and recalls of contaminated food[88] Identification of association between oropharyngeal cancer and human papillomavirus infection has led to new prevention efforts[89]

Table 1. (Continued.)

Data Source	Strengths	Weaknesses	Examples of Effects on Policy or Practice
Cross-sectional studies	Provide snapshot of exposure and outcome Can help generate hypotheses Can be conducted rapidly	Difficult to attribute causality Difficult to control for confounding	Evaluation of association between sodium intake and blood pressure, along with other evidence, has provided support for policy interventions aimed at reducing sodium consumption[46,48] Evaluation of deep venous thrombosis in hospitals led to the identification of a low rate of use of appropriate preventive measures and to improved practices[90]
Ecologic studies	Provide population-level vs. individual-level data Can document outcomes of natural experiments Can be conducted rapidly	Potential for invalid conclusions from noncausal associations because of residual confounding (ecologic fallacy) Potential for data that are not standardized or comparable	Vaccine-effectiveness studies have led to changes in immunization recommendations, increasing the proportion of people protected[29] Analysis of mortality in heat waves resulted in practical recommendations to mitigate weather-related effects[91]
Pragmatic trials and large observational studies	Potential for high generalizability Can be conducted at relatively low cost Potential to emulate real-world experience in application of findings, increasing external validity	Potential for varying quality of data Potential for adoption of some interventions by control group, biasing results toward a null result Potential for increased likelihood of invalid results because of a lack of standardization of assessment, treatment, and adherence Potential for loss to follow-up to affect interpretation of results	Study comparing treatments for type 2 diabetes was 20 times larger and had much longer follow-up than previous randomized, controlled trials, resulting in clear evidence for clinical decision making[63] and fewer patients being treated with a sulfonylurea, a drug class not previously known to be associated with increased mortality Trial provided evidence that task sharing among nurses and other health workers did not reduce quality of care for patients with human immunodeficiency virus infection[92]
Program-based evidence	May provide definitive evidence of efficacy in real-world conditions	Without control community, may not be possible to determine causality Potential for control community to adopt some of the interventions, biasing results toward a null result	New Zealand Back to Sleep campaign provided definitive evidence that advice given to parents about having babies sleep in a supine position could prompt actions that would reduce the incidence of SIDS,[9] leading to global programs that have greatly reduced infant mortality Implementation of public health measures such as tobacco taxes, smoke-free laws, and educational campaigns have documented efficacy in ways that would not have been possible or definitive otherwise and has led to widespread implementation of tobacco-control measures that save millions of lives[93]
Case reports and series	Can provide inexpensive, detailed assessments Useful for evaluation of rare diseases, identification of rare events Can lead to reasonable conclusions about relative benefit of different treatments for rare diseases	Limited ability to draw definitive conclusions because of the lack of a comparison group Selection bias (e.g., patients with rapid resolution or rapid progression to death may be underrepresented)	Identification of the acquired immunodeficiency syndrome and other newly recognized conditions (e.g., Zika virus–associated microcephaly and newly identified drug-resistant organisms or mechanisms) has accelerated improvements in detection, treatment, and prevention of these conditions[94,95] Highly effective treatments have been identified for conditions that otherwise had poor prognoses (e.g., penicillin as a broad-spectrum antibiotic)[96]
Registries	Determine efficacy in real life Can provide useful data for rare diseases Can help assess quality of care Can provide results rapidly	Difficult or impossible to control for confounding and bias	Studies have documented and improved quality of care and determined the most effective treatment of patients undergoing dialysis, reducing the incidence of preventable complications and deaths[97] Studies have determined predictors of survival in pulmonary arterial hypertension, enabling more informed treatment choices[98]

potentially highly valuable sources of data is counterproductive. A better approach is to clarify the health outcome being sought and determine whether existing data are available that can be rigorously and objectively evaluated, independently of or in comparison with data from RCTs, or whether new studies (RCT or otherwise) are needed.

New ways of obtaining valuable health data continue to emerge. "Big data," including information from electronic health records and expanded patient registries, along with increased willingness of patients to participate and share health information, are generating useful data for large interventional studies and providing new opportunities for complementary use of multiple data sources to gain stronger evidence for action.[71] For example, although an RCT may show the benefit of a drug, large observational studies can be conducted to refine dosages and identify rare adverse events. In addition, new strategies have been undertaken to increase the efficacy and efficiency of RCTs, including collaborative and adaptive trials to increase enrollment, reduce costs and time to completion, and better identify populations that benefit from treatments.[72-74] Advances in genomic science may allow for better understanding of unique characteristics in patients that can affect outcomes of RCTs and other studies and be used to improve the validity of study findings.

There is no single, best approach to the study of health interventions; clinical and public health decisions are almost always made with imperfect data (Table 1). Promoting transparency in study methods, ensuring standardized data collection for key outcomes, and using new approaches to improve data synthesis are critical steps in the interpretation of findings and in the identification of data for action, and it must be recognized that conclusions may change over time. There will always be an argument for more research and for better data, but waiting for more data is often an implicit decision not to act or to act on the basis of past practice rather than best available evidence. The goal must be actionable data — data that are sufficient for clinical and public health action that have been derived openly and objectively and that enable us to say, "Here's what we recommend and why."

The views expressed in this article are those of the author and do not necessarily represent the views of the Centers for Disease Control and Prevention or the Department of Health and Human Services.

Disclosure forms provided by the author are available with the full text of this article at NEJM.org.

I thank Kathryn Foti, M.P.H., Drew Blakeman, M.S., and Robin Moseley, M.A.T., for assistance with preparation of an earlier version of the manuscript and review of relevant literature, and Joanna Taliano, M.L.S., for assistance with literature searches.

REFERENCES

1. Centers for Disease Control and Prevention. Mission, role and pledge (http://www.cdc.gov/about/organization/mission.htm).
2. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?" Lancet 2005;365:82-93.
3. Jones DS, Podolsky SH. The history and fate of the gold standard. Lancet 2015;385:1502-3.
4. Bothwell LE, Greene JA, Podolsky SH, Jones DS. Assessing the gold standard — lessons from the history of RCTs. N Engl J Med 2016;374:2175-81.
5. Kones R, Rumana U, Merino J. Exclusion of 'nonRCT evidence' in guidelines for chronic diseases — is it always appropriate? The Look AHEAD study. Curr Med Res Opin 2014;30:2009-19.
6. Chavez-MacGregor M, Giordano SH. Randomized clinical trials and observational studies: is there a battle? J Clin Oncol 2016;34:772-3.
7. Deaton A. Instruments, randomization, and learning about development. J Econ Lit 2010;48:424-55.
8. Woodcock J, Ware JH, Miller PW, McMurray JJV, Harrington DP, Drazen JM. Clinical trial series. N Engl J Med 2016;374:2167.
9. Mitchell EA, Scragg R, Stewart AW, et al. Results from the first year of the New Zealand Cot Death Study. N Z Med J 1991;104:71-6.
10. Irving M, Eramudugolla R, Cherbuin N, Anstey KJ. A critical review of grading systems: implications for public health policy. Eval Health Prof 2016 May 10 (Epub ahead of print).
11. Sacks H, Chalmers TC, Smith H Jr. Randomized versus historical controls for clinical trials. Am J Med 1982;72:233-40.
12. Jackson LA, Jackson ML, Nelson JC, Neuzil KM, Weiss NS. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. Int J Epidemiol 2006;35:337-44.
13. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med 2000;342:1887-92.
14. Golfam M, Beall R, Brehaut J, et al. Comparing alternative design options for chronic disease prevention interventions. Eur J Clin Invest 2015;45:87-99.
15. Grodstein F, Stampfer MJ, Manson JE, et al. Postmenopausal estrogen and progestin use and the risk of cardiovascular disease. N Engl J Med 1996;335:453-61.
16. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. JAMA 2002;288:321-33.
17. Prentice RL, Langer R, Stefanick ML, et al. Combined postmenopausal hormone therapy and cardiovascular disease: toward resolving the discrepancy between observational studies and the Women's Health Initiative clinical trial. Am J Epidemiol 2005;162:404-14.
18. Prentice RL, Langer RD, Stefanick ML, et al. Combined analysis of Women's

Health Initiative observational and clinical trial data on postmenopausal hormone treatment and cardiovascular disease. Am J Epidemiol 2006;163:589-99.
19. Manson JE, Chlebowski RT, Stefanick ML, et al. Menopausal hormone therapy and health outcomes during the intervention and extended poststopping phases of the Women's Health Initiative randomized trials. JAMA 2013;310:1353-68.
20. Vandenbroucke JP. The HRT controversy: observational studies and RCTs fall in line. Lancet 2009;373:1233-5.
21. Mendelsohn ME, Karas RH. HRT and the young at heart. N Engl J Med 2007; 356:2639-41.
22. Nasal influenza vaccine approved for younger children. Food and Drug Administration Consumer Updates. September 26, 2007 (https://www.fda.gov/ForConsumers/ConsumerUpdates/ucm048715.htm).
23. Ashkenazi S, Vertruyen A, Arístegui J, et al. Superior relative efficacy of live attenuated influenza vaccine compared with inactivated influenza vaccine in young children with recurrent respiratory tract infections. Pediatr Infect Dis J 2006; 25:870-9.
24. Fleming DM, Crovari P, Wahn U, et al. Comparison of the efficacy and safety of live attenuated cold-adapted influenza vaccine, trivalent, with trivalent inactivated influenza virus vaccine in children and adolescents with asthma. Pediatr Infect Dis J 2006;25:860-9.
25. Belshe RB, Edwards KM, Vesikari T, et al. Live attenuated versus inactivated influenza vaccine in infants and young children. N Engl J Med 2007;356:685-96.
26. Grohskopf LA, Olsen SJ, Sokolow LZ, et al. Prevention and control of seasonal influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices (ACIP) — United States, 2014–15 influenza season. MMWR Morb Mortal Wkly Rep 2014;63:691-7.
27. Gaglani M, Pruszynski J, Murthy K, et al. Influenza vaccine effectiveness against the 2009 Pandemic A (H1N1) virus differed by vaccine-type during 2013-14 in the United States. J Infect Dis 2016;213: 1546-56.
28. Flannery B. Update on effectiveness of live-attenuated versus inactivated influenza vaccines in children and adolescents aged 2-18 years – US Flu VE Network. Presented at the Meeting of the Advisory Committee on Immunization Practices, Atlanta, October 29–30, 2014 (http://www.nitag-resource.org/uploads/media/default/0001/02/a4c56839e81370e3d8635468dddee24ac681d2da.pdf).
29. Chung JR, Flannery B, Thompson MG, et al. Seasonal effectiveness of live attenuated and inactivated influenza vaccine. Pediatrics 2016;137(2):e20153279.
30. ACIP votes down use of LAIV for 2016-2017 flu season. Media statement from the Centers for Disease Control and Prevention, June 22, 2016 (http://www.cdc.gov/media/releases/2016/s0622-laiv-flu.html).
31. Streptomycin in Tuberculosis Trials Committee. Streptomycin treatment of pulmonary tuberculosis. Br Med J 1948;2: 769-82.
32. Iseman MD, Sbarbaro JA. Short-course chemotherapy of tuberculosis: hail Britannia (and friends)! Am Rev Respir Dis 1991;143:697-8.
33. Mitchison DA. The diagnosis and therapy of tuberculosis during the past 100 years. Am J Respir Crit Care Med 2005;171:699-706.
34. Fox W. Self-administration of medicaments: a review of published work and a study of the problems. Bull Int Union Tuberc 1962;32:307-31.
35. Dawson JJY, Devadatta S, Fox W, et al. A 5-year study of patients with pulmonary tuberculosis in a concurrent comparison of home and sanatorium treatment for one year with isoniazid plus PAS. Bull World Health Organ 1966;34:533-51.
36. Treatment of tuberculosis: guidelines for national programmes. Geneva: World Health Organization, 1997.
37. Zwarenstein M, Schoeman JH, Vundule C, Lombard CJ, Tatley M. Randomised controlled trial of self-supervised and directly observed treatment of tuberculosis. Lancet 1998;352:1340-3.
38. Akkslip S, Rasmithat S, Maher D, Sawert H. Direct observation of tuberculosis treatment by supervised family members in Yasothorn Province, Thailand. Int J Tuberc Lung Dis 1999;3:1061-5.
39. Sterling TR, Lehmann HP, Frieden TR. Impact of DOTS compared with DOTS-plus on multidrug resistant tuberculosis and tuberculosis deaths: decision analysis. BMJ 2003;326:574.
40. Moonan PK, Quitugua TN, Pogoda JM, et al. Does directly observed therapy (DOT) reduce drug resistant tuberculosis? BMC Public Health 2011;11:19.
41. Frieden TR, Sherman LF, Maw KL, et al. A multi-institutional outbreak of highly drug-resistant tuberculosis: epidemiology and clinical outcomes. JAMA 1996; 276:1229-35.
42. Xu JQ, Murphy SL, Kochanek KD, Bastian BA. Deaths: final data for 2013. Natl Vital Stat Rep 2016;64(2):1-119.
43. Nwankwo T, Yoon SS, Burt V, Gu Q. Hypertension among adults in the United States: National Health and Nutrition Examination Survey, 2011-2012. NCHS Data Brief 2013;133:1-8.
44. Institute of Medicine. Strategies to reduce sodium intake in the United States. Washington, DC: National Academies Press, 2010.
45. Mozaffarian D, Fahimi S, Singh GM, et al. Global sodium consumption and death from cardiovascular causes. N Engl J Med 2014;371:624-34.
46. He FJ, Pombo-Rodrigues S, Macgregor GA. Salt reduction in England from 2003 to 2011: its relationship to blood pressure, stroke and ischaemic heart disease mortality. BMJ Open 2014; 4(4):e004549.
47. He FJ, MacGregor GA. Salt reduction lowers cardiovascular risk: meta-analysis of outcome trials. Lancet 2011;378:380-2.
48. Cook NR, Appel LJ, Whelton PK. Sodium intake and all-cause mortality over 20 years in the trials of hypertension prevention. J Am Coll Cardiol 2016;68:1609-17.
49. Mente A, O'Donnell M, Rangarajan S, et al. Associations of urinary sodium excretion with cardiovascular events in individuals with and without hypertension: a pooled analysis of data from four studies. Lancet 2016;388:465-75.
50. Graudal N, Jürgens G, Baslund B, Alderman MH. Compared with usual sodium intake, low- and excessive-sodium diets are associated with increased mortality: a meta-analysis. Am J Hypertens 2014;27:1129-37.
51. O'Donnell M, Mente A, Rangarajan S, et al. Urinary sodium and potassium excretion, mortality, and cardiovascular events. N Engl J Med 2014;371:612-23.
52. Cobb LK, Anderson CA, Elliott P, et al. Methodological issues in cohort studies that relate sodium intake to cardiovascular disease outcomes: a science advisory from the American Heart Association. Circulation 2014;129:1173-86.
53. Willett W. Nutritional epidemiology: monographs in epidemiology and biostatistics. 3rd ed. New York: Oxford University Press, 2013.
54. He FJ, MacGregor GA. Hypertension: salt: flawed research should not divert actions to reduce intake. Nat Rev Nephrol 2016;12:514-5.
55. Cogswell ME, Maalouf J, Elliott P, Loria CM, Patel S, Bowman BA. Use of urine biomarkers to assess sodium intake: challenges and opportunities. Annu Rev Nutr 2015;35:349-87.
56. Rakova N, Jüttner K, Dahlmann A, et al. Long-term space flight simulation reveals infradian rhythmicity in human Na(+) balance. Cell Metab 2013;17:125-31.
57. Cogswell ME, Mugavero K, Bowman BA, Frieden TR. Dietary sodium and cardiovascular disease risk — measurement matters. N Engl J Med 2016;375:580-6.
58. Multiple Risk Factor Intervention Trial Group. The Multiple Risk Factor Inter-

vention Trial (MRFIT): a national study of primary prevention of coronary heart disease. JAMA 1976;235:825-7.
59. Frieden TR. Sodium reduction — saving lives by putting choice into consumers' hands. JAMA 2016;316:579-80.
60. Schieppati A, Henter J-I, Daina E, Aperia A. Why rare diseases are an important medical and social issue. Lancet 2008;371:2039-41.
61. Griggs RC, Batshaw M, Dunkle M, et al. Clinical research for rare disease: opportunities, challenges, and solutions. Mol Genet Metab 2009;96:20-6.
62. Rubinstein YR, Groft SC, Bartek R, et al. Creating a global rare disease patient registry linked to a rare diseases biorepository database: Rare Disease-HUB (RD-HUB). Contemp Clin Trials 2010;31:394-404.
63. Frakt AB. An observational study goes where randomized clinical trials have not. JAMA 2015;313:1091-2.
64. Lederle FA, Cushman WC, Ferguson RE, Brophy MT, Fiore Md LD. Chlorthalidone versus hydrochlorothiazide: a new kind of Veterans Affairs cooperative study. Ann Intern Med 2016;165:663-4.
65. Ernst ME, Lund BC. Renewed interest in chlorthalidone: evidence from the Veterans Health Administration. J Clin Hypertens (Greenwich) 2010;12:927-34.
66. Roush GC, Holford TR, Guddati AK. Chlorthalidone compared with hydrochlorothiazide in reducing cardiovascular events: systematic review and network meta-analyses. Hypertension 2012;59:1110-7.
67. ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). JAMA 2002;288:2981-97.
68. Dorsch MP, Gillespie BW, Erickson SR, Bleske BE, Weder AB. Chlorthalidone reduces cardiovascular events compared with hydrochlorothiazide: a retrospective cohort analysis. Hypertension 2011;57:689-94.
69. Woolever DR. The art and science of clinical decision making. Fam Pract Manag 2008;15:31-6.
70. Groopman J. How doctors think. Boston: Houghton Mifflin, 2007.
71. Angus DC. Fusing randomized trials with big data: the key to self-learning health care systems? JAMA 2015;314:767-8.
72. Thall PF, Wathen JK. Practical Bayesian adaptive randomisation in clinical trials. Eur J Cancer 2007;43:859-66.
73. Moss AJ, Francis CW, Ryan D. Collaborative clinical trials. N Engl J Med 2011;364:789-91.
74. Bhatt DL, Mehta C. Adaptive designs for clinical trials. N Engl J Med 2016;375:65-74.
75. Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. Lancet 2002;360:7-22.
76. Sever PS, Dahlöf B, Poulter NR, et al. Prevention of coronary and stroke events with atorvastatin in hypertensive patients who have average or lower-than-average cholesterol concentrations, in the Anglo-Scandinavian Cardiac Outcomes Trial — Lipid Lowering Arm (ASCOT-LLA): a multicentre randomised controlled trial. Lancet 2003;361:1149-58.
77. Colhoun HM, Betteridge DJ, Durrington PN, et al. Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS): multicentre randomised placebo-controlled trial. Lancet 2004;364:685-96.
78. Khan M, Qureshi AI. Factors associated with increased rates of post-procedural stroke or death following carotid artery stent placement: a systematic review. J Vasc Interv Neurol 2014;7:11-20.
79. Baigent C, Keech A, Kearney PM, et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. Lancet 2005;366:1267-78.
80. Cholesterol Treatment Trialists' (CTT) Collaboration. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet 2010;376:1670-81.
81. Schrag D, Kuntz KM, Garber JE, Weeks JC. Decision analysis — effects of prophylactic mastectomy and oophorectomy on life expectancy among women with *BRCA1* or *BRCA2* mutations. N Engl J Med 1997;336:1465-71.
82. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, Wilschut J, van Ballegooijen M, Kuntz KM. Evaluating test strategies for colorectal cancer screening: a decision analysis for the U.S. Preventive Services Task Force. Ann Intern Med 2008;149:659-69.
83. Hayes JH, Ollendorf DA, Pearson SD, et al. Active surveillance compared with initial treatment for men with low-risk prostate cancer: a decision analysis. JAMA 2010;304:2373-80.
84. Tai V, Grey A, Bolland MJ. Results of observational studies: analysis of findings from the Nurses' Health Study. PLoS One 2014;9(10):e110403.
85. Niederau C, Lange S, Heintges T, et al. Prognosis of chronic hepatitis C: results of a large, prospective cohort study. Hepatology 1998;28:1687-95.
86. Zou B, Li T, Zhou Q, et al. Adjuvant therapeutic modalities in primary small cell carcinoma of esophagus patients: a retrospective cohort study of multicenter clinical outcomes. Medicine (Baltimore) 2016;95(17):e3507.
87. Mulrooney DA, Yeazel MW, Kawashima T, et al. Cardiac outcomes in a cohort of adult survivors of childhood and adolescent cancer: retrospective analysis of the Childhood Cancer Survivor Study cohort. BMJ 2009;339:b4606.
88. Gottlieb SL, Newbern EC, Griffin PM, et al. Multistate outbreak of listeriosis linked to turkey deli meat and subsequent changes in US regulatory policy. Clin Infect Dis 2006;42:29-36.
89. D'Souza G, Kreimer AR, Viscidi R, et al. Case–control study of human papillomavirus and oropharyngeal cancer. N Engl J Med 2007;356:1944-56.
90. Cohen AT, Tapson VF, Bergmann J-F, et al. Venous thromboembolism risk and prophylaxis in the acute hospital care setting (ENDORSE study): a multinational cross-sectional study. Lancet 2008;371:387-94.
91. Mastrangelo G, Fedeli U, Visentin C, Milan G, Fadda E, Spolaore P. Pattern and determinants of hospitalization during heat waves: an ecologic study. BMC Public Health 2007;7:200.
92. Fairall L, Bachmann MO, Lombard C, et al. Task shifting of antiretroviral treatment from doctors to primary-care nurses in South Africa (STRETCH): a pragmatic, parallel, cluster-randomised trial. Lancet 2012;380:889-98.
93. Wisotzky M, Albuquerque M, Pechacek TF, Park BZ. The National Tobacco Control Program: focusing on policy to broaden impact. Public Health Rep 2004;119:303-10.
94. Pneumocystis pneumonia — Los Angeles. MMWR Morb Mortal Wkly Rep 1981;30:250-2.
95. Schuler-Faccini L, Ribeiro EM, Feitosa IM, et al. Possible association between Zika virus infection and microcephaly — Brazil, 2015. MMWR Morb Mortal Wkly Rep 2016;65:59-62.
96. Fleming A. On the antibacterial action of cultures of a Penicillium, with special reference to their use in the isolation of B. influenzae. Br J Exp Pathol 1929;10:226-36.
97. Shroff GR, Frederick PD, Herzog CA. Renal failure and acute myocardial infarction: clinical characteristics in patients with advanced chronic kidney disease, on dialysis, and without chronic kidney disease: a collaborative project of the United

States Renal Data System/National Institutes of Health and the National Registry of Myocardial Infarction. Am Heart J 2012;163:399-406.

98. Benza RL, Miller DP, Gomberg-Maitland M, et al. Predicting survival in pulmonary arterial hypertension: insights from the Registry to Evaluate Early and Long-Term Pulmonary Arterial Hypertension Disease Management (REVEAL). Circulation 2010;122:164-72.