

The Ethics of Testing a Test: Randomized Trials of the Health Impact of Diagnostic Tests for Infectious Diseases

David W. Dowdy,^{1,3} Celine R. Gounder,^{3,4,a} Elizabeth L. Corbett,^{6,7} Lucky G. Ngwira,⁷ Richard E. Chaisson,^{1,2,3,4} and Maria W. Merritt^{2,5}

Departments of ¹Epidemiology, and ²International Health, Johns Hopkins Bloomberg School of Public Health, ³Center for Tuberculosis Research, ⁴Division of Infectious Diseases, Department of Medicine, and ⁵Berman Institute of Bioethics, Johns Hopkins University, Baltimore, Maryland; ⁶Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, United Kingdom; and ⁷Malawi-Liverpool-Wellcome Trust Research Programme, Blantyre

In the last decade, many new rapid diagnostic tests for infectious diseases have been developed. In general, these new tests are developed with the intent to optimize feasibility and population health, not accuracy alone. However, unlike drugs or vaccines, diagnostic tests are evaluated and licensed on the basis of accuracy, not health impact (eg, reduced morbidity or mortality). Thus, these tests are sometimes recommended or scaled up for purposes of improving population health without randomized evidence that they do so. We highlight the importance of randomized trials to evaluate the health impact of novel diagnostics and note that such trials raise distinctive ethical challenges of equipoise, equity, and informed consent. We discuss the distinction between equipoise for patient-important outcomes versus diagnostic accuracy, the equity implications of evaluating health impact of diagnostics under routine conditions, and the importance of offering reasonable choices for informed consent in diagnostic trials.

Central to the control of most infectious diseases is appropriate and timely diagnosis. In the last decade, numerous new rapid diagnostic tests for human immunodeficiency virus (HIV; eg, point-of-care HIV antibody assays and CD4 cell counts [1]), malaria (eg, rapid diagnostic tests [RDTs] [2]), tuberculosis (eg, Xpert MTB/RIF, a fully automated molecular test for active tuberculosis [3, 4]), and other infectious diseases have been developed. In general, these new tests are developed with the intent to optimize feasibility and population health, not accuracy alone. Often, reference-

standard tests exist (eg, laboratory-based CD4 cell counts, high-quality microscopy for malaria, tuberculosis culture) that are more accurate or provide important ancillary information (eg, speciation, complete drug susceptibility testing) but are not deployable in resource-limited settings and may be difficult to maintain at high quality in low-burden settings. The novel tests are not intended to supplant the current reference standard but rather to make diagnosis more feasible and thereby improve population health.

This reality—wherein health impact is prioritized over laboratory-based accuracy—has important ethical implications. Unlike drugs or vaccines, diagnostic tests are designed “to indicate the presence or absence of disease [5],” not to prevent or treat disease per se. Thus, diagnostics are—and should be—evaluated in a fundamentally different fashion than biological agents. Drugs and vaccines improve health directly and must be evaluated in vivo, whereas diagnostics depend on a larger process to improve population health and are evaluated outside the body. As a result, unlike drugs

Received 19 June 2012; accepted 16 August 2012; electronically published 23 August 2012.

^aPresent affiliation: Bureau of Tuberculosis Control, New York City Department of Health and Mental Hygiene, New York.

Correspondence: David W. Dowdy, MD, PhD, 615 N Wolfe St, E6531, Baltimore, MD 21205 (ddowdy@jhsph.edu).

Clinical Infectious Diseases 2012;55(11):1522–6

© The Author 2012. Published by Oxford University Press on behalf of the Infectious Diseases Society of America. All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com.

DOI: 10.1093/cid/cis736

or vaccines, diagnostics are commonly licensed without randomized evidence that their use will improve health. For example, whereas a drug that killed plasmodia in vitro would not be licensed without randomized evidence that it reduced morbidity in vivo, malaria diagnostics (RDTs) are routinely licensed without evidence that their desirable diagnostic characteristics translate into improved health. Furthermore, diagnosis involves not just sensitivity and specificity, but rather a cascade of processes from care seeking by the patient through delivery of appropriate treatment [6]. Novel diagnostic tests may appropriately aim to improve many different aspects of this cascade, including accuracy, cost, and patient satisfaction, but changing one process in the cascade (eg, accuracy) may have complex and even paradoxical effects on other interrelated processes (eg, provider treatment decisions). Differential standards for licensure are not per se ethically problematic: biological agents and diagnostic tests serve different purposes, and their evaluation should reflect this distinction. Difficulties arise, however, when novel diagnostics, once licensed, are widely recommended for scale-up—with the explicit goal of improving population health—before obtaining randomized evidence that their use will accomplish that goal [7]. A basic ethical requirement for responsible public health practice is to evaluate the assumption that a proposed program achieves its goals of improving population health. If data do not exist to “demonstrate the program’s assumptions,” then “ethically, the program should not be implemented” [8].

Proof of a test’s diagnostic accuracy in an optimal laboratory setting should not be taken as proof that use of the test will be effective in reducing population morbidity or mortality. For example, greater sensitivity in detecting prostate cancer (by analogy from diagnosis of chronic disease) may increase morbidity without reducing mortality [9], and tuberculosis culture (a sensitive test) often has minimal impact on clinical decision making in high-burden settings [10]. Licensure of these tests is not ethically problematic, but scaling them up with the intent to improve population health would be. Indeed, in some cases, the use of new high-sensitivity diagnostics may actually erode patient confidence and worsen system performance (eg, by disrupting established algorithms, increasing service downtime, or providing untreatable diagnoses). If patients and health professionals lose confidence in routine health services, high-sensitivity tests are of little value. Thus, although initial trials of diagnostic accuracy may demonstrate that a novel diagnostic is more sensitive than the standard of care, subsequent randomized trials of health outcomes are needed before scaling up those diagnostics for the purpose of improving health.

Cluster randomized trials (CRTs) randomize participants in clusters (eg, attendees of a clinic that is implementing a new diagnostic test) and evaluate outcomes (eg, disease morbidity

or mortality) on the cluster level. Unlike drugs or vaccines, where a placebo and active agent can be delivered in the same manner on an individual level, deployment of a new diagnostic test requires changes in the entire diagnostic process (eg, infrastructure, staff training, specimen flow), making individual randomization infeasible. Moreover, although convenience is generally not appropriate justification for trial design, individual randomization (with person-by-person selection of a new vs old test) is often both infeasible and unrepresentative of actual practice in settings with constrained resources that must maintain a simple, high-throughput, “public health” approach for diagnosis of common illnesses. For all of these reasons, cluster randomization is often the ideal type of trial design to evaluate the health impact of using a new diagnostic test [11]. CRTs raise specific ethical concerns [12]. Here, we raise 3 distinctive ethical challenges associated with CRTs of new diagnostic tests for infectious diseases that take health (eg, morbidity or mortality) as their primary outcome (Figure 1).

THREE ETHICAL CHALLENGES: EQUIPOISE, EQUITY, AND INFORMED CONSENT

Equipoise

Effectiveness trials comparing the current standard of care in resource-limited settings to a new diagnostic test with known superior accuracy (as measured in an optimal laboratory setting) may appear ethically troubling to some. Investigators knowingly direct the control arm to offer a less accurate diagnostic test, apparently contrary to the requirement that clinical research offer each individual participant (or cluster of participants) a favorable net risk-benefit ratio [13–15]. Said another way, the trial appears to lack *clinical equipoise*, understood as “uncertainty or disagreement within the expert clinical community about the relative merits” of the intervention and the control arms [16, 17]. Clinical equipoise is widely (although not universally) held to be “a necessary ethical condition for the commencement of a trial [13],” and its existence has been debated fiercely, for example, with the development of the Xpert MTB/RIF test for tuberculosis [18, 19].

This troubling appearance, however, results from applying the concept of clinical equipoise to the comparative *accuracy* of the 2 diagnostic *tests* rather than the comparative *health impact* of the 2 diagnostic *processes*. Much as certainty about a drug’s proximal effect (eg, the ability of folate to reduce homocysteine levels) is compatible with uncertainty about the drug’s health impact (reduction in cardiovascular events [20]), certainty about a diagnostic test’s superior accuracy is compatible with uncertainty about whether its associated process will reduce morbidity or mortality. In both cases, surrogate end points should not replace clinical end points without evidence that an intervention’s ability to improve the surrogate end

1. **Equipose**
Challenge: Knowledge of superior diagnostic accuracy does not imply absence of clinical equipoise regarding impact on health.

Solution: Randomized trials should demonstrate improved health outcomes (morbidity and mortality) before implementing public health programs (eg, scale-up of diagnostic tests) that target such outcomes.
2. **Equity**
Challenge: The impact of diagnostic tests is setting specific, and failure to implement diagnostics in the proper setting (eg, central versus peripheral health center) carries greatest risk to the underserved.

Solution: Randomized trials of novel diagnostic tests should be carried out in multiple settings, under routine conditions that can reasonably be replicated during scale-up.
3. **Informed Consent**
Challenge: In cluster-randomized trials of novel diagnostic tests, diagnostic processes (eg, index of suspicion, empiric treatment patterns) are changed by adoption of the novel test.

Solution: Informed consent documents should appropriately present the risks of declining diagnosis with the novel test, including the possibility that the new test may change the diagnostic process even for persons who decline participation.

Figure 1. Randomized trials of the health impact of diagnostic tests for infectious diseases: ethical challenges and proposed solutions.

point (diagnostic accuracy) will result in an improvement in the clinical end point (population health). When this relationship is not clearly demonstrated, randomized trials of health outcomes are ethically justified. More accurate techniques (eg, microscopic observation drug susceptibility assay for tuberculosis [21]) may require retraining, development of new testing algorithms, and additional maintenance resulting in less reliable results when resources (eg, trained staff) are limited. More-sensitive tests can result in low-specificity algorithms, causing harm through false-positive results and diverted resources (eg, tuberculosis serology [22]) or by increasing drug prescriptions to patients without disease (eg, malaria RDTs, in some circumstances [23, 24]). In the case of diagnostic testing, certainty about diagnostic accuracy does not dissolve the state of clinical equipoise with respect to health outcomes, nor render randomized trials of those outcomes unnecessary.

Equity

CRTs evaluating the health impact of new diagnostic tests, especially in resource-limited settings, also raise ethical challenges regarding equity. The efficacy of a drug regimen is often generalizable across settings; it may depend on local resistance patterns but rarely on local infrastructure or disease prevalence. By contrast, the health impact of a diagnostic test depends highly on geographic setting (eg, ambient temperature and humidity), disease prevalence, and health system factors including infrastructure, specimen-transport chains, and ability to service diagnostic machinery. It is therefore not

sufficient to demonstrate the efficacy of a diagnostic test in a single setting; to evaluate the health impact of novel diagnostics among the disadvantaged populations whom they are designed to benefit, studies must employ pragmatic designs in many alternative settings, using resources likely to be available during posttrial scale-up. These studies will often employ non-experimental designs, but randomized trials are frequently required to demonstrate that diagnostic tests result in improved health outcomes.

For example, unlike therapeutic trials, diagnostic effectiveness trials may require separate evaluations of the same test at different levels of the health system. Different diagnostic processes and resources may be available at centralized (eg, reference laboratory) versus peripheral (eg, health clinic) health facilities during scale-up, requiring different trials to evaluate them before scale-up occurs. Performing a single trial at the centralized level might appear more ethically palatable from the perspective of producing replicable scientific results, but runs the risk of exacerbating inequity. In the absence of comparative evidence on health impact from both centralized and peripheral levels, local policy makers could either restrict implementation to tertiary facilities where laboratory-based accuracy data can best be replicated or decentralize implementation without a proven health benefit. Restricting implementation to centralized facilities runs the risk of excluding disadvantaged populations from possible benefit, while decentralized implementation runs the risk of diverting scarce resources designated for the poor without improving their health. In either event, the poor and

disadvantaged would be disproportionately affected: implementation of the very diagnostics intended to improve their health ultimately has the perverse outcome of exacerbating inequity. Trials of health outcomes under diverse programmatic conditions promote equity by showing policy makers how to scale up novel diagnostics in a way that maximizes health benefit for disadvantaged populations.

Informed Consent

As others have noted, informed consent for randomization may not be required if “it is not possible to approach subjects at the time of randomization” (as may happen in CRTs) [25]. However, CRTs of novel diagnostics in resource-limited settings present additional challenges for informed consent, particularly for individuals attending facilities randomized to the novel diagnostic under study (the intervention arm). Novel diagnostic tests may alter clinical practice patterns, such that individuals in the intervention arm who decline the new test may still receive care that is influenced by the new test’s presence. For example, the availability of a novel test may increase clinicians’ index of suspicion, make therapeutic action after a positive result more likely [26], and change clinical thresholds [27] for empiric treatment or obtaining additional diagnostic information. Thus, a new diagnostic test should be seen as a “package” incorporating an associated process, potentially altering patients’ entire clinical experience and precluding the choice to receive the previous standard of care as an alternative to the novel diagnostic.

From the perspective of informed consent, then, individuals attending intervention facilities have a choice to accept or decline the novel diagnostic test; if they decline, the previous standard of care is often not an option. Given that prospective ethical review of any trial ought to assure “a favorable net risk-benefit ratio” (including minimization of risk) to participants [15], the risk-benefit ratio of accepting the novel diagnostic should be (and generally is) favorable as compared with declining it. Accordingly, an accurate informed consent process should not always frame this choice as a neutral one. Rather, it should provide information about the novel diagnostic test and process, its attendant risks and prospect of benefit, and the risks of declining the test, including the possibility that the presence of the new test may have changed existing diagnostic processes. Suggesting to patients in the intervention arm of a diagnostic trial that they have access to the same diagnostic process as was present before the trial is generally misleading.

CONCLUSION

A central issue in infectious disease control over the next 5 years will be the appropriate deployment of novel diagnostic tests

with known accuracy but uncertain health impact. Randomized trials of the use of diagnostics in resource-limited settings raise distinctive ethical challenges of equipoise, equity, and informed consent that are not well addressed under our current ethical and regulatory framework. The design and ethical review of these trials should incorporate 3 ethical considerations (1):

- Equipoise regarding the health outcomes of a novel diagnostic process can coexist with certainty about the diagnostic accuracy of the novel test it incorporates.
- Equity requires that the health impact of diagnostic tests be evaluated under routine conditions with the resources and processes expected to be available during scale-up.
- In CRTs of novel diagnostics, the informed consent process in the intervention arm should acknowledge that the alternative to trial participation may not be the previous standard of care.

To reduce the tremendous global burden of infectious disease morbidity and mortality, we must continue to develop novel diagnostics explicitly designed to benefit populations in high-burden, resource-limited areas. However, we can reconcile our enthusiasm to implement these diagnostics with our obligation to improve population health only by demonstrating their impact on patient-important outcomes in routine settings. Forthrightly addressing these distinctive challenges of equipoise, equity, and informed consent will allow us to proceed in improving diagnosis while maintaining firm ethical footing.

Notes

Financial support. This work was supported by the US National Institutes of Health (grants R01 AI093316-01A1 and 1P30AI094189-01A1), the Bill & Melinda Gates Foundation (grant 19790.01 for the Consortium to Respond Effectively to the AIDS-TB Epidemic), and the Greenwall Faculty Scholars Program in Bioethics. The funding sources had no role in study design, data collection and analysis, the decision to publish, or the preparation of the manuscript.

Potential conflicts of interest. D. W. D., C. R. G., E. L. C., L. G. N., and R. E. C. are conducting an ongoing cluster randomized trial of Xpert MTB/RIF for intensified case finding of tuberculosis in a resource-limited setting. M. W. M. reports no potential conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Jani IV, Siteo NE, Alfai ER, et al. Effect of point-of-care CD4 cell count tests on retention of patients and rates of antiretroviral therapy initiation in primary health clinics: an observational cohort study. *Lancet* 2011; 378:1572–9.
2. Abba K, Deeks JJ, Olliaro P, et al. Rapid diagnostic tests for diagnosing uncomplicated *P. falciparum* malaria in endemic countries. *Cochrane Database Syst Rev* 2011; 7:CD008122.
3. Boehme CC, Nicol MP, Nabeta P, et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet* 2011; 377:1495–505.

4. Boehme CC, Nabeta P, Hillemann D, et al. Rapid molecular detection of tuberculosis and rifampin resistance. *N Engl J Med* **2010**; 363:1005–15.
5. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* **1978**; 299:926–30.
6. Dowdy DW, Cattamanchi A, Steingart KR, Pai M. Is scale-up worth it? challenges in economic analysis of diagnostic tests for tuberculosis. *PLoS Medicine* **2011**; 8:e1001063.
7. World Health Organization. Guidelines for the treatment of malaria. 2nd ed. Geneva, Switzerland: World Health Organization, **2010**.
8. Kass NE. An ethics framework for public health. *Am J Public Health* **2001**; 91:1776–82.
9. Andriole GL, Crawford ED, Grubb RL 3rd, et al. Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med* **2009**; 360:1310–9.
10. Stall N, Rubin T, Michael JS, et al. Does solid culture for tuberculosis influence clinical decision making in India? *Int J Tuberc Lung Dis* **2011**; 15:641–6.
11. Donner A, Klar N. Design and analysis of cluster randomization trials in health research. London, United Kingdom: Arnold, **2000**.
12. Weijer C, Grimshaw JM, Taljaard M, et al. Ethical issues posed by cluster randomized trials in health research. *Trials* **2011**; 12:100.
13. Binik A, Weijer C, McRae AD, et al. Does clinical equipoise apply to cluster randomized trials in health research? *Trials* **2011**; 12:118.
14. Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? *JAMA* **2000**; 283:2701–11.
15. Emanuel EJ, Wendler D, Grady C. An ethical framework for biomedical research. In: Emanuel EJ, Grady C, Crouch RA, Lie RK, Miller FG, Wendler D, eds. *The Oxford textbook of clinical research ethics*. New York, NY: Oxford University Press, **2008**.
16. Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med* **1987**; 317:141–5.
17. Joffe S, Truog R. Equipoise and randomization. In: Emanuel EJ, Grady C, Crouch RA, Lie RK, Miller FG, Wendler D, eds. *The Oxford textbook of clinical research ethics*. New York NY: Oxford University Press, **2008**:245–60.
18. Trebucq A, Enarson DA, Chiang CY, et al. Xpert® MTB/RIF for national tuberculosis programmes in low-income countries: when, where and how? *Int J Tuberc Lung Dis* **2011**; 15:1567–72.
19. Singh JA, Bhan A. The ethics of national tuberculosis programmes in low-income countries not rolling out Xpert® MTB/RIF. *Int J Tuberc Lung Dis* **2011**; 15:1563.
20. Bona KH, Njolstad I, Ueland PM, et al. Homocysteine lowering and cardiovascular events after acute myocardial infarction. *N Engl J Med* **2006**; 354:1578–88.
21. Moore DAJ, Evans CAW, Gilman RH, et al. Microscopic-observation drug-susceptibility assay for the diagnosis of TB. *N Engl J Med* **2006**; 355:1539–50.
22. Dowdy DW, Steingart KR, Pai M. Serological testing versus other strategies for diagnosis of active tuberculosis in India: a cost-effectiveness analysis. *PLoS Med* **2011**; 8:e1001074.
23. Hamer DH, Ndhlovu M, Zurovac D, et al. Improved diagnostic testing and malaria treatment practices in Zambia. *JAMA* **2007**; 297:2227–31.
24. Uzochukwu BS, Onwujekwe E, Ezuma NN, Ezeoke OP, Ajuba MO, Sibeudu FT. Improving rational treatment of malaria: perceptions and influence of RDTs on prescribing behaviour of health workers in southeast Nigeria. *PLoS One* **2011**; 6:e14627.
25. McRae AD, Weijer C, Binik A, et al. When is informed consent required in cluster randomized trials in health research? *Trials* **2011**; 12:202.
26. Davis JL, Katamba A, Vasquez J, et al. Evaluating tuberculosis case detection via real-time monitoring of tuberculosis diagnostic services. *Am J Respir Crit Care Med* **2011**; 184:362–7.
27. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med* **1980**; 302:1109–17.