AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Clinical Microbiology®

COMMENTARY

Evaluating Novel Diagnostics in an Outbreak Setting: Lessons Learned from Ebola

Nira R. Pollock,[a] Betsy Wonderly[b]

Department of Laboratory Medicine, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, USA[a]; Foundation for Innovative New Diagnostics, Geneva, Switzerland[b]

ABSTRACT Inadequate access to rapid testing for Ebola virus disease during the 2014-to-2016 outbreak led to an explosion in the development of diagnostics that could be performed at or near the point of care and by less-experienced operators, leading in turn to an acute need for novel test evaluation. Here, we present the challenges to development and evaluation of novel diagnostics in an emergency setting and suggestions for potential new "global emergency standards" to address them.

KEYWORDS Ebola, diagnostic, point of care, outbreak, diagnostics, Ebola virus

Accepted manuscript posted online 1 March 2017

Citation Pollock NR, Wonderly B. 2017. Evaluating novel diagnostics in an outbreak setting: lessons learned from Ebola. J Clin Microbiol 55:1255–1261. https://doi.org/10.1128/JCM.00053-17.

Editor Colleen Suzanne Kraft, Emory University

Address correspondence to Nira R. Pollock, nira.pollock@childrens.harvard.edu.

The views expressed in this Commentary do not necessarily reflect the views of the journal or of ASM.

The challenges to laboratory diagnosis of Ebola virus disease (EVD) during the 2014-to-2016 outbreak in West Africa were substantial and have been well documented (1, 2). Testing, if available, relied primarily on standard high-complexity real-time reverse transcription-PCR (RT-PCR) performed in biosafety level 4 (BSL-4) biocontainment laboratories, and operational challenges inherent to every step of the testing process (venous blood collection, sample transport and testing, and result reporting) led to diagnostic errors and substantial delays in the return of results. Inadequate access to rapid testing for EVD and persistent operational and infrastructural challenges led to an unprecedented explosion in the development of diagnostic tests that could be performed at or near the point of care and by less-experienced operators; this development in turn led to an acute need for evaluation of these novel diagnostic products. We were deeply involved in efforts to evaluate novel EVD diagnostics in Sierra Leone: N.P. led field studies to evaluate (i) the performance of the ReEBOV Antigen Rapid Test kit (Corgenix, Inc.) for both point-of-care and laboratory-based testing (3) and (ii) the laboratory-based performance of the GeneXpert Ebola assay (Cepheid, Inc.) (4), and B.W. led the diagnostic team from Foundation for Innovative New Diagnostics (FIND) that worked collaboratively with the WHO and others (3) on test evaluations, including a comparative assessment of the novel lateral flow immunoassays for EVD diagnosis (5; B. Wonderly et al., unpublished data). Through this work, we experienced firsthand both the substantial international drive toward effective collaboration around rapid development and evaluation of diagnostics and the systematic challenges to evaluating novel diagnostics in an emergency setting. Without substantial collective consideration of these specific challenges and identification of their solutions, development and evaluation of diagnostics in future outbreak scenarios will be similarly handicapped, again blunting our ability to respond rapidly and save lives. Here, we present an outline of these challenges to rapid evaluation of diagnostics and suggestions for potential new global emergency standards that could be developed to address them.

At the start of the 2014-to-2016 EVD outbreak, there were no EVD diagnostics with either FDA or WHO approval. As the critical international laboratory deployment effort roared into action, multiple laboratory-developed tests (including, for example, tests developed by the U.S. Centers for Disease Control, the U.S. Department of Defense, and

Public Health England) and one commercial product (altona Diagnostics, Hamburg, Germany) were implemented throughout West Africa. Despite the urgency of the outbreak, there was insufficient data sharing regarding the design and performance of these assays, making it difficult to know how tests compared and which test might be an optimal benchmark for novel test evaluation. For those preparing field evaluations of the numerous novel diagnostics in development by both commercial and academic parties, it was difficult to pinpoint which novel tests to prioritize for study and which test to use as a reference method, and it was difficult to avoid actual or perceived bias in test selection.

Navigating the sea of formal and informal processes for study approvals in the affected countries was also tremendously challenging, leading to both confusion and delay. Not only was it difficult to understand which governmental or nongovernmental body was in charge of each element of the process (e.g., human subject protocol review and registration of products under evaluation), but even upon identification of the responsible agency (e.g. see reference 6), it was a challenge to clarify the exact approval process. Often, procedures that were in place did not anticipate the unique requirements of handling hemorrhagic fever viruses or obtaining emergency authorizations. It was also difficult to access and secure clinical samples (e.g., blood, buccal swabs, etc.) to include in test evaluations. Standards and regulations around patient consent and sample ownership were undefined and research priorities uncertain.

The lack of clarity around specimens, processes, and both local and global priorities, in addition to the unique biosafety considerations for sample collection and testing for EVD, led to substantial delays in the overall diagnostic evaluation process. Tragically, in the midst of this delay, desperation for rapid diagnostic test access on the ground led to clinical use of insufficiently validated tests. Lack of collaboration and communication between stakeholders attempting to evaluate novel test platforms led to redundancy, inefficiency, and even bias in test evaluation data. And even after novel diagnostics that made it through the FDA and WHO emergency use authorization processes (which themselves were insufficiently aligned) were validated in studies in the affected countries, the confusion around the processes required for in-country regulatory approval and the lack of clear algorithms for test usage delayed or prevented tests from actually being used.

Table 1 presents a detailed list of specific challenges and questions encountered during the evaluation of novel EVD diagnostics during the 2014-to-2016 outbreak. The first category of questions concerns sample ownership. Many groups pursuing test development and evaluation were in search of clinical samples essential to those activities, but there was considerable confusion about ownership of those samples once available clinical testing (RT-PCR) had been performed. Ownership was particularly unclear in a context in which field labs operated by many different countries and organizations were testing samples from patients across West Africa, some of whom were themselves travelers within the region. Lack of clarity about sample ownership confused processes for access to samples both within country (for those wishing to evaluate tests in the field) and internationally (for those pursuing novel test development in other countries). It was unclear in the chaos of the outbreak who should be in charge of making decisions about any payments for samples (particularly for commercial test developers), sharing of samples, and research priorities for sample use, and uncertainty continues even in the aftermath of the outbreak regarding appropriate use of precious stored samples.

A second category of questions concerns data ownership. Similar to ownership of the samples, there was confusion concerning ownership of the clinical data generated during routine clinical testing and whether those data could be managed via cloud-based servers or had to be communicated and stored locally. Ultimately, email (i.e., Gmail) was used to communicate much of the clinical and research data, leveraging cloud-based servers. An additional challenge was the dissemination of research findings. Given lengthy publication timelines and potential barriers to access to published articles, informal distribution of data was done locally during the outbreak to ensure that the information could be used right away to improve clinical care and impact

TABLE 1 Specific challenges and questions encountered during evaluation of novel EVD diagnostics during the 2014-to-2016 outbreak[a]

Specific challenge	Questions encountered	Possible response(s)/follow-up question(s)
Sample ownership	Who owns clinical samples, once clinical test results have been generated?	The government of the country where patients were tested (e.g., SL, for testing performed there)? The government of the country whose patients were tested (e.g., Guinea, for Guinean patients tested in SL)? The government of the country directing the lab in which patient samples were tested (e.g., UK, operating in SL)? The organization responsible for testing and storing patient samples (e.g., PHE, operating in SL)? The organization funding the laboratory testing and/or sample storage (e.g., DFID)? The WHO or other global governing body? The patient?
	Can samples be shipped out of the country of origin for test development (or research use) elsewhere? Should developers of diagnostic and therapeutic solutions have to pay for access to clinical samples? Who decides if—and with whom—samples should be shared? Who determines research priorities?	
	What efforts receive priority?	Academic research? Commercial development of new tests?
Data ownership	Who owns data that have been generated as a result of clinical testing? (e.g., clinical results, operational data)	The government of the country where patients were tested (e.g., SL, for testing performed there)? The government of the country whose patients were tested, generating the data (e.g., Guinea, for Guinean patients tested in SL)? The government of the country directing the lab in which patient samples were tested, generating the data (e.g., UK, operating in SL)? The organization responsible for testing and storing patient samples and generating and storing the data (e.g., PHE, operating in SL)? The organization funding the laboratory testing samples and generating the data (e.g., DFID)? The WHO or other global governing body?
	Can data, particularly clinical data, be transmitted out of the country of origin for test development, optimization of disease management, or research use elsewhere during an outbreak? Should data consumers (WHO, test developers, etc.) have to pay for data access? Who decides if—and with whom—data should be shared? Who pays for data transmission, where applicable, and what system(s) should be used for transmission of clinical data?	
	What efforts receive priority?	Disease management? Academic research? Commercial development of new tests, therapies, and/or vaccines?
Human subjects	Do patients need to provide consent for research testing done on their own excess clinical samples?	If "yes," logistically, how would this be carried out? Consider the following factors: Clinical condition/competency for consent Lack of literacy/education/understanding Language/cultural barriers
	Which bodies within a country need to sign off on a human subjects research protocol?	Formal vs informal (courtesy) notifications? For example, in SL: MoH, LTWG, Pharmacy Board?
Regulatory authority	Regarding WHO EUAL vs FDA EUA, will countries accept either or neither as sufficient for in-country use of the product?	

(Continued on next page)

TABLE 1 (Continued)

Specific challenge	Questions encountered	Possible response(s)/follow-up question(s)
	Will countries require additional in-country data to be generated prior to in-country approval for use of the product?	
	Which group, within a given country, is legally responsible for clearing a test for clinical use?	For example, in SL: LTWG or Pharmacy Board, or both? Role of MoH?
	How do manufacturers validate their system when access to samples is limited and biosafety concerns complicate testing?	
	Who should be responsible for organizing/facilitating test validation?	Manufacturers? CDC? FIND? WHO? Others? (If not the manufacturer, who decides which tests to prioritize for evaluation?)
	Which approval (e.g., WHO EUAL, FDA EUA, or CE-Marking) determines which sample types can be used for diagnostic testing in country?	Can/should a clinical lab validate an "unapproved" sample type and then report results for clinical use?
Identifying a "gold standard" reference technology	How do we approach LDTs that are not available commercially?	Should an LDT be considered as a candidate reference method? What comparative evaluations (vs commercial assays or vs other LDTs) should be required to allow utilization of an LDT as a reference method? Should efforts be made to make an LDT widely available during an outbreak? Who should pay for distribution of an LDT?
	How might determination of cutoff thresholds (e.g., C_T values) be standardized and/or publicized to better allow interlab comparisons of assays and results?	
Effective communication	How do we effectively communicate research findings, product approvals, testing algorithm updates, guideline updates, biosafety concerns, etc., to clinicians and programs working in the field, where infrastructure may be lacking?	

[a]Abbreviations: EVD, Ebola virus disease; SL, Sierra Leone; UK, United Kingdom; PHE, Public Health England; DFID, Department for International Development; WHO, World Health Organization; MOH, Ministry of Health; LTWG, Laboratory Technical Working Group; EUAL, emergency use authorization and listing; EUA, emergency use authorization; CDC, Centers for Disease Control and Prevention; FIND, Foundation for Innovative New Diagnostics; FDA, Food and Drug Administration; CE-Marking, European conformity marking; LDT, laboratory-developed test; C_T, cycle threshold.

disease transmission. Finally, new questions around access to clinical data arose toward the end of the outbreak, when novel diagnostics were newly capable of automatic reporting in real time. Given the potential economic impact of a positive test result as the outbreak waned, national decision makers preferred to be able to review the data with local leadership prior to sharing it more widely.

Questions about the participation of human subjects in the test development and evaluation process included both questions about provision of consent for use of excess clinical samples (typically solved by waiver of informed consent) and about which bodies within a given country actually needed to approve a human subject research protocol. Protocol implementation required coordination of submission to multiple committees for review, including in-country investigational review boards (IRBs), the IRBs of the organizations for each of the participating research groups (which often included academic centers/organizations in one country and a collaborating laboratory from another country), and potentially review committees representing the lab processing the clinical samples and the funders supporting the labs and the Ebola treatment centers. In addition to submission of the IRB protocols themselves, it was necessary to navigate (without direct guidance) processes for courtesy notification of additional groups within the country who had a stake in study implementation or outcome.

Questions around regulatory authority hindered both the implementation of protocols for test development/evaluation and the implementation of the tests themselves

after successful evaluation. While the FDA and WHO made great efforts to quickly develop emergency use authorizations (EUAs) as tests became available, it was clear neither what clearance for clinical use those EUAs actually conferred in those nations affected by the epidemic nor whether each country required generation of additional data in country to allow a test to be used there for clinical decision making. Furthermore, the in-country processes for obtaining clearance of a test for clinical use were not clear to either those providing clinical care within the country (and wishing to use new tests that seemed to perform well) or those wishing to help make new tests available for use (e.g., FIND); this confusion also extended to which sample types (fingerstick, buccal swab, venipuncture blood, etc.) were actually approved for use in each country. Overall, it was not clear who should actually take the lead in organizing and facilitating test evaluations, leading to less effective cooperation between manufacturers, academics, nongovernmental organizations, and governing bodies.

Given that evaluations of novel diagnostics always require comparison of the novel test to existing reference methods, decisions about which reference method(s) should be considered the "gold" (reference) standard for test evaluation are critical. Early in the Ebola outbreak, the fact that only one commercially available RT-PCR kit (altona) was available and that many of the newly established field laboratories brought their own laboratory-developed tests into the field caused substantial confusion regarding which test might be suitable or optimal as a reference method for field evaluations of novel tests. Limited data on the performance of individual tests, whether laboratory developed or commercial, was available to test developers and researchers to shed light on the comparative performance of assays in use in the field. Furthermore, limited access to field laboratories and samples effectively required "selection" of reference methods based on availability rather than on knowledge of test performance. The altona assay was viewed by many (including the WHO) to be an attractive reference method in the early stages of the outbreak, given its commercial availability. However, the assay was found to perform less well than expected when compared to a laboratory-developed test (3), potentially due to suboptimal implementation under field laboratory conditions. Even if laboratory-developed tests in use by field labs did have optimal performance, test developers and evaluation researchers often did not have access to those tests to consider their use as a reference method. Finally, even data generated under the auspices of formal test evaluation for regulatory clearance were often presented without sufficient detail (e.g., sample handling, processing, and cutoff values) to allow optimal interstudy comparisons of either reference method or novel test performance.

The final category of challenges relates to the need for effective communication in order not only to evaluate novel tests, but also to move those with successful evaluation results into clinical use and to move updated testing guidelines and recommendations to the field. Overall, it was challenging to (i) effectively communicate research findings to those in a position to make decisions about in-country approvals, (ii) get products approved, and (iii) move approved products into clinical use using appropriate testing algorithms.

Table 2 presents specific suggestions for the development of "global emergency standards" to address the challenges and questions posed in Table 1; such standards could perhaps be developed by an expert committee led by the WHO, with international input. First, a systematic framework should be developed for diagnostic evaluations during WHO-declared "global emergencies," so that field labs, test developers, and clinical sites have templates to work with to expedite and streamline protocol development for evaluation of both laboratory-based and point-of-care technologies. Second, standards should be developed that summarize required IRB approvals and required pathways to those approvals (including both formal and informal approvals needed), along with clear IRB template documents.

TABLE 2 Moving toward global emergency standards for diagnostic evaluation in an outbreak setting[a]

Specific goal	Method(s) to achieve goal
Develop systematic framework for diagnostic evaluations	Provide templates (for customization) to field labs, test developers, and clinical sites Flexibly accommodate test development/evaluation work already in progress at time of emergency declaration
Develop standards summarizing required IRB approvals and pathways to approval	Share publicly and provide to manufacturers and groups preparing for test evaluations In-country IRBs should provide protocol templates for guidance Clarify formal vs informal approvals needed (and order of approval) Anticipate/accommodate different types of evaluations (discarded samples, POC testing)
Optimize collaboration between global leaders (WHO and CDC/FDA)	Eliminate redundant efforts; synergize Avoid bias regarding which tests to evaluate Facilitate efficiency and synergy in test evaluations (e.g., multiple tests run in parallel on a given sample and provision of standardized samples for external quality assurance) Develop an effective communication plan for relaying new developments to clinicans and programs in the field, acknowledging infrastructure limitations (limited wifi and/or cellular coverage)
Increase transparency	Develop standards for data sharing, including in advance of publication If not publishing, increase level of detail in EUA documentation to allow critical analysis and test comparison

[a]The overall goal is to develop a common understanding and approach or, more formally, global standards around sample/data ownership and test validation/regulation during WHO-declared "global emergencies." Abbreviations: WHO, World Health Organization; EUA, emergency use authorization; CDC, Centers for Disease Control and Prevention; FDA, Food and Drug Administration; IRB, investigational review board; POC, point of care.

Third, it is critical that collaborations between global leaders (WHO and CDC/FDA) be optimized. This includes harmonization of EUA procedures, elimination of redundant work, and promotion of synergy (e.g., collaborative work to allow multiple tests to be performed in parallel on a single sample and provision of standardized samples for external quality assurance), avoidance of bias and subjectivity in selection of tests for evaluation, and development of an effective communication plan for relaying new developments to country leads, clinicians, and programs in the field. Fourth, transparency must be increased, through both development of clear standards for data sharing in advance of publication and a requirement for increased level of detail in EUA documentation to allow critical analysis of data and test comparison. With a concerted collaborative effort, each of the steps proposed in Table 2 could be established in advance of the next outbreak. Additionally, proactive work (including collaborative biobank development) could facilitate advance development and validation of diagnostics for emerging pathogens of known global outbreak concern.

Recent efforts by the WHO toward development of an "R&D blueprint for action to prevent epidemics" (7, 8) as a road map to accelerate development and evaluation of therapeutics, vaccines, and diagnostics for pathogens with outbreak potential are encouraging and hopefully will ultimately address the specific challenges encountered in our diagnostic evaluation experience. While the current draft is focused on development of therapeutics and vaccines, elements of the WHO R&D blueprint that will address funding, data sharing, biobanking, and regulatory approval are just as relevant to diagnostics development and evaluation. As stated in the draft blueprint (7), the entire global community will benefit if development efforts during outbreaks are "facilitated through adoption of fair and transparent principles which will have been negotiated by all stakeholders ahead of an emergency." We must work together now to develop such principles and standards for development and evaluation of diagnostics in an outbreak setting, so that the lessons learned from our Ebola experience are not wasted.

ACKNOWLEDGMENTS

We thank Bill Rodriguez, Jana Broadhurst, Timothy Brooks, Madhukar Pai, and Megan Murray for helpful comments on the manuscript.

N.R.P. and B.W. have no conflicts of interest.

REFERENCES

1. Broadhurst MJ, Brooks TJ, Pollock NR. 2016. Diagnosis of Ebola virus disease: past, present, and future. Clin Microbiol Rev 29:773–793. https://doi.org/10.1128/CMR.00003-16.
2. Perkins MD, Kessel M. 2015. What Ebola tells us about outbreak diagnostic readiness. Nat Biotechnol 33:464–469. https://doi.org/10.1038/nbt.3215.
3. Broadhurst MJ, Kelly JD, Miller A, Semper A, Bailey D, Groppelli E, Simpson A, Brooks T, Hula S, Nyoni W, Sankoh AB, Kanu S, Jalloh A, Ton Q, Sarchet N, George P, Perkins MD, Wonderly B, Murray M, Pollock NR. 2015. ReEBOV Antigen Rapid Test kit for point-of-care and laboratory-based testing for Ebola virus disease: a field validation study. Lancet 386:867–874. https://doi.org/10.1016/S0140-6736(15)61042-X.
4. Semper AE, Broadhurst MJ, Richards J, Foster GM, Simpson AJ, Logue CH, Kelly JD, Miller A, Brooks TJ, Murray M, Pollock NR. 2016. Performance of the GeneXpert Ebola assay for diagnosis of Ebola virus disease in Sierra Leone: a field evaluation study. PLoS Med 13:e1001980. https://doi.org/10.1371/journal.pmed.1001980.
5. World Health Organization. WHO emergency use authorization and listing for EVD IVDs. Public report. http://www.who.int/diagnostics_laboratory/procurement/150908_final_public_report_ea_0021-020-00.pdf.
6. Office for Human Research Protections, US Department of Health and Human Services. 2016. International compilation of human research standards. Accessed October 27, 2016.
7. World Health Organization. 2017. An R&D blueprint for action to prevent epidemics: plan of action. http://www.who.int/csr/research-and-development/WHO-R_D-Final10.pdf?ua=1. Accessed 16 February 2017.
8. World Health Organization. 2017. An R&D blueprint for action to prevent epidemics: progress update, February 2017. http://www.who.int/csr/research-and-development/BP_Feb_update_1022017.pdf?ua=1. Accessed 16 February 2017.