Tolerable tests

Citation for published version:
https://doi.org/10.1177/01622439241252709

Digital Object Identifier (DOI):
10.1177/01622439241252709

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published in:
Science, Technology, & Human Values (ST&HV)

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Alice Street¹ and Ann H. Kelly²

Abstract
The response to the 2014-2016 West African epidemic was a watershed for emergency research and innovation, forcing a shift in regulatory norms as evidentiary standards were pitted against humanitarian imperatives and biosecurity concerns. This article examines how those ethical and epistemic negotiations unfolded in practice through the development, testing, and use of novel tools for Ebola diagnosis with a focus on Sierra Leone. We track the priorities placed on the accuracy, feasibility, and clinical efficiency of Ebola diagnostic platforms and explore how these varied over the course of the outbreak and for different actors involved in their deployment. The lack of clarity over which tools might be fit for purpose exposed the profound ambiguities around the nature, scope, and purpose of building in-country Ebola diagnostic capacity. Ultimately, we argue that the accelerated

¹School of Social and Political Science, The University of Edinburgh, United Kingdom
²Department of Global Health, School of Global Affairs, Faculty of Social Science & Public Policy, King’s College London

Corresponding Author:
Alice Street, School of Social and Political Science, The University of Edinburgh, 15a George Square, Edinburgh EH8 9LD, United Kingdom. Email: alice.street@ed.ac.uk
regulatory process coordinated by the World Health Organization operated as a liminal procedure that both revealed the scientific, ethical, and political trade-offs and inequalities attendant to an emerging regime of emergency research and development, and provided a tentative, reflexive platform for regulatory experimentation, deliberation, and reform.

**Keywords**
Ebola, diagnostics, regulatory science, emergency research and development, tolerability

**Introduction**
On July 7, 2015, eleven months after the declaration of the West Africa Ebola outbreak as a Public Health Emergency of International Concern (PHEIC), the World Health Organization (WHO) launched the Emergency Use Assessment and Listing (EUAL) procedure (WHO 2015a). Described as a “time-limited procedure” for the assessment of quality, safety, and efficacy/performance of medical products during an outbreak based on a “minimal level of information” (WHO 2015a), the EUAL provided a market prompt and regulatory pathway to catalyze the development of medical countermeasures (drugs, vaccines, and diagnostics). The emergency procedure substantially diluted the evidentiary requirements of the WHO’s usual prequalification process for medical products; however, considering the lack of products at a mature stage in the product life-cycle and the magnitude of what was fast becoming a humanitarian crisis, allowances had to be made (see Gostin and Friedman 2014). In a public health emergency, the WHO argued, “the community may be willing to tolerate less certainty about the performance and safety of products, given the morbidity and/or mortality of the disease and the shortfall of diagnostic options” (WHO 2015a, 1). The purpose of the mechanism, then, was not to provide conclusive assurance of a novel product’s quality and efficacy. Instead, the EUAL was designed to identify thresholds of tolerability for promising, though yet unproven, tools (drugs, vaccines, and diagnostics) in a context in which radical uncertainty was accepted as a feature of rapid outbreak response.

The regulation of biomedical innovations is hardly straightforward in the best of circumstances. As a science, regulation is fundamentally pragmatic, pulling together heterodox strands of evidence and expert opinion
“mandated” in response to clinical need (Salter 1998). That process of adjudication is rife with controversy. Beyond the pressures and perils of approving potentially lifesaving or life-threatening innovations for public use, regulatory decisions are meted out within a vast and highly lucrative medical market and thus are vulnerable to conflict of interest and industry capture (McGoey 2007; Mulinar and Davis 2020). Regulatory science always operates in the wild (Callon, Lascoumes, and Barthe 2011), its epistemological judgments inflected by explicitly ethical and value-laden concerns (Funtowicz and Ravetz 1993; Fortun and Fortun 2005). Operating between “facts” and “values,” “science,” and “politics,” regulation derives credibility not from the purity of its evidence but from the rigor of its process: the degree to which those decisions can be felt to adhere to bureaucratic conventions and localized cultures of expertise (Jasanoff 1990).

During a PHEIC, the stakes of regulatory decisions gain considerably in amplitude and uncertainties around “tolerability” proliferate far beyond the questions of safety and accuracy posed by the WHO upon the launch of the EUAL. When a deadly outbreak spills across national boundaries, between and across health systems, encountering gross inequities, the value of medical technologies becomes bound to the speed and facility with which they can be deployed. The deliberate and iterative assemblage of evidence through which regulatory institutions build credibility and distance from political and commercial interests are attenuated by the exigencies of demand and limits in supply. That demand is, moreover, highly dynamic: the changing epidemiological profile of a disease can quickly alter what tools are needed and where, further torquing any efforts to produce a formalized risk-based or precautionary calculus, while historically embedded geographic inequalities (Benton 2015; Hirsch 2020) circumscribe the very processes for how diagnostic needs are determined and the “thresholds of tolerability” for answering those needs.

As intertwined uncertainties around diagnostic quality, accuracy, rapidity, efficacy, and ethics unfolded during the Ebola outbreak, a novel cluster of regulatory challenges emerged for the international community: how to accelerate innovation while maintaining quality and reliability; how to meet immediate needs while respecting national sovereignty; how to determine those needs, and the appropriate trade-offs that might be necessary to meet them in a rapidly changing epidemiological landscape. Emergency innovations require oversight that is pluralistic and polycentric (see Hogarth, Miller, and Sturdy 2022, 2); regulatory decisions must be brokered across the structures of multiple legal and technical authorities and between diverse communities of practice. The makeshift international architecture
established to oversee diagnostic regulation during the Ebola outbreak is, we argue, best characterized by its liminality, a term Sethi (2018, 127) uses “to conceptualize medical innovation as a space which breaks down pre-existing structures,” with the “potential for transformation across different thresholds.” This framing of medical innovation as processual transboundary work is particularly well suited to describe the role of the WHO as a quasi-regulatory, international body, and the tools that it deployed in response to the particular ambiguities that characterize emergency research and development (R&D), including the EUAL mechanism, the Target Product Profile (TPP), and field validation protocols. Liminality also speaks to the normative dislocation of innovations deployed to address health emergencies that cut across traditional sovereign boundaries, transcending the authoritative reach of national and regional regulatory bodies (Demortain 2017).

By tracking the regulatory trajectories of intravenous diagnostic tests that passed through the EUAL procedure during the 2014-2016 West Africa Ebola outbreak, this article plots a course through the thicket of scientific and ethical “liminalities” that accompany the regulation of medical products in public health emergencies. Our focus remains on Sierra Leone, the worst-affected country in the region—a public health context marked by years of intensive, if punctuated, foreign intervention—and the chosen location for performance evaluations of EUAL candidate tests (Benton 2015; Erikson 2019; Leach and Fairhead 2008; Wilkinson and Fairhead 2017). We start by describing the process through which needs for diagnostics were articulated and, subsequently, the complexity of determining the regulatory thresholds through which those needs could feasibly be met. Following individual tests into the field, we describe how those benchmarks became unmoored by the epidemiological dynamism of the outbreak but also by the governmental fault lines between the WHO’s technical and moral authority and the legal authority of National Regulatory Authorities (NRAs).

In our discussion, we interrogate the politics of global health R&D equity, attending to the priorities of different actors and the degree to which they could be brought into alignment within the emergency regulatory space. We find that while hastily assembled, highly uneven, and fragile, the EUAL provided a sufficiently supple regulatory scaffold for diverse stakeholders to collectively adjudicate trade-offs in evidence and speed, quality, and access against which a workable emergency response could be mounted. By the very nature of its time-limited and provisional evocation, the EUAL created a context to articulate the emergency as a series of
situated issues and reflexively grasp the value of novel technologies in providing ways to act upon them (cf., Bruckermann 2023; Millo and Lezaun 2006). To conclude, we consider the degree to which that liminality might be extended to form the basis for more deliberative, equitable, and open-ended frameworks for global health innovation in the future.

In pursuing this line of inquiry, our article builds upon long-term ethnographic engagement with the diagnostics sector, including more than eighty interviews with product developers, public health experts, emergency responders, local patients, and medical professionals in the United States, Europe, and Sierra Leone. It is also informed by Kelly’s experience of working as part of the 2014-2016 West African Ebola response and continued policy and research engagement with research teams, scientific and ethics advisory bodies, including the WHO. We have contextualized those engagements by conducting reviews of gray and scholarly literature regarding diagnostic development and emergency response, including that generated by the WHO, ministries of health, and other global health organizations.

At the heart of this article is a concern with the generative possibilities of this experimental or “liminal space” of regulation, and what new global health norms, standards, roles, and responsibilities the emergent practices of emergency R&D might precipitate. By approaching the EUAL as a liminal process, through which what is tolerable can be “worked out” (Fortun and Fortun 2005, 50; see also Nading 2015), we ask in what ways may liminal regulation be reflexively embraced and extended to form the basis of more deliberative, and sufficiently supple, frameworks for global health innovation in the future.²

**Diagnostic Aspirations**

Diagnostic tests rapidly took center stage in the international response to the Ebola outbreak in West Africa. Prompt diagnosis was vital to isolating symptomatic cases and rapidly tracing contacts in the community. Delays in diagnosis threatened the health of patients in holding units and increased the risk of the disease spreading further in the population (WHO 2015b). Quick, postmortem diagnosis permitted the release of bodies for burial (a significant point of tension with local communities, for whom appropriate funerary rites were a priority, see Lipton 2017) and was essential for promptly establishing whether it was safe to release survivors back into the community (Broadhurst, Brooks, and Pollock 2016; Kost et al. 2015; Kost 2018). Diagnostics were also seen to play a critical role in restoring public
trust in the response and countering widespread reluctance to access health services during the epidemic. It was the WHO’s (2014a) hope that “having such tests readily available could restore some order to West African health systems, which have been devastated by fear of contagion as well as by the demands of managing a deadly and dreaded disease.” By linking orders of knowledge and action, and carving out specific problems and sites for intervention, diagnostics rendered the outbreak “doable” in critical ways (Fujimura 1987).

At the onset of the outbreak, the region’s laboratory infrastructure was chronically under-resourced, beset by decades of colonial neglect, a de-prioritization of diagnostics in international health policy, and vertical health programs that failed to invest in generalized laboratory capacity (Vernooij 2021; Koster et al. 2021; Wilkinson 2017). While international aid agencies and research groups rapidly established an ad hoc network of mobile and field laboratories staffed by international laboratory workers and equipped with machines, devices, and reagents shipped in from overseas, these laboratories remained uncoordinated, highly centralized, and unevenly distributed, necessitating extensive transportation networks to move samples, patients, and results (Vernooij 2021). In-house manual reverse transcription-polymerase chain reaction (RT-PCR) techniques for nucleic acid testing were “cumbersome, slow, complex and costly” and lacked standardization and external validation (WHO 2014b). This disorganized and inefficient beginning to the diagnostic response led to a devastating loss of life in the early weeks (Chertow et al. 2014) and prompted the WHO to reflect that “diagnostics had been lost in the scramble for Ebola control” (WHO and FIND 2014, 3).

Against the backdrop of a laboratory system characterized by disconnection and delay, the commercial manufacture of rapid, portable diagnostic devices offered an attractive route to standardizing, speeding up, and scaling up tests. The expectations placed on diagnostics rapidly cohered around two new rapid testing technologies: highly accurate automated RT-PCR nucleic acid testing platforms, which could provide results in a few hours; and highly mobile and easy-to-use lateral flow assays, commonly referred to as rapid diagnostic tests (RDTs), which could provide less accurate results in a matter of minutes. Over the preceding decade, the rollout of similar testing devices for other diseases had transformed the delivery of infectious disease programs (Chandler et al. 2011; Engel et al. 2017; Engel et al. 2022; Harper and Khatri 2019). Now, Ebola presented a test case for the accelerated development and deployment of these technologies in an emergency outbreak scenario for which these technologies promised to
significantly amplify the rapidity and scale of response (see also Kameda et al. 2021).

Pledging to do everything it could “to minimize the barriers faced by diagnostic companies to develop and deploy their tests” (WHO 2014a, see Figure 1 also 2014c), the WHO (2014b) published a “target product profile for Zaire ebolavirus rapid, simple test”—essentially a list of performance characteristics that developers should aim for in new diagnostic tools. The list included both “acceptable” and “desired” characteristics—an aspirational account of what diagnostics could and should involve. Tacking between the optimal characteristics of the lateral flow assay (mobility, speed) and those of the automated RT-PCR platform (accuracy), the WHO’s TPP implicitly laid out two primary testing scenarios for rapid tests in the outbreak. Lateral flow assays, WHO officials hoped, could be used as screening devices to support initial quarantine decisions in rural communities without access to laboratory infrastructure; RT-PCR platforms would provide highly accurate confirmatory tests in more centralized Ebola Treatment Units. While never used as a formal criterion for listing, the TPP was nonetheless viewed as a key reference point for manufacturers seeking to submit a product to the EUAL. Sitting somewhere between market device and evaluative instrument (Street 2018), the TPP might be considered a kind of liminal regulatory device; a tool that marks out a field of acceptable action between the possible and the aspirational, without defining any firm thresholds for action or carrying any legal weight. The aspirational nature of the TPP would gain and lose salience as the development process unfolded and as the realities of the Ebola response took hold.

**Evidentiary Thresholds**

Soon after the release of the TPP, the WHO invited manufacturers to begin submitting their products for “emergency assessment,” a process that would later evolve into the formal EUAL procedure. The ambiguous status of the EUAL as a process of assessment but not authorization spoke to the WHO’s unique status within the architecture of global health governance, which imbues it with the moral authority to set standards and the technical authority to provide guidance for National Regulatory Agencies but accords it no legal authority of its own. Thus, rather than providing legal authorization for the use of a medical product within any specific country, the EUAL procedure conferred a multilateral endorsement for the procurement of lifesaving tools through United Nations (UN) and affiliated agencies. By relaxing evidentiary thresholds and facilitating access to humanitarian
markets, the WHO hoped that the EUAL would mobilize investment in the rapid development of new medical products that the biotech industry otherwise deemed too complex and unprofitable to pursue.4

The evaluation of manufacturer submissions to the WHO’s emergency assessment mechanism for diagnostics entailed a three-step process. First, the submission of a dossier outlining the firm’s quality management systems and manufacturing capacity. Second, submission of product performance specifications and any available validation and verification studies. Third, products considered to meet suitable thresholds of quality,

**Figure 1.** Target product profile for Zaire ebolavirus rapid, simple test to be used in the control of the Ebola outbreak in West Africa (WHO 2014b).
manufacturing capacity, and performance were subjected to an independent laboratory-based performance evaluation study using clinical samples. WHO’s established prequalification team took a lead in this process, but it was supported by the Ebola Diagnostic Action Committee, a coordinating group set up by the WHO, Doctors Without Borders (MSF), and FIND, a donor-funded foundation focused on supporting diagnostic innovation for global health. The respective involvement of these organizations varied at different stages of the process and over the course of the outbreak. Where WHO personnel initially ran the in-country performance evaluations, for example, they soon found their capacity to keep up with submissions to be limited, and this aspect of the process became delegated to FIND. FIND also took the lead in mapping the existing landscape for Ebola IVDs and identifying those manufacturers most likely to be able to commercialize their products within the anticipated time frame of the PHEIC.

A major difficulty for the prequalification team was that manufacturers did not always, as one interviewee put it, “submit enough information for the WHO to be able to assess the claims.” Because manufacturers were asked to provide whatever information they had, the submissions were not standardized. More fundamentally, the benchmarks for diagnostic quality for an aspirational product were hazy, so bringing those submissions into comparative alignment posed a problem. Critical shortages in blood samples (a problem of access we detail below) placed further ethical pressure on decisions about which manufacturers to select to go forward for further laboratory evaluation. As the WHO’s diagnostics lead explained to us, “We didn’t want to give samples to technology that had no chance of being developed, that was too early.”

In light of these evidentiary gaps and evaluative bottlenecks, manufacturing capacity—the question of whether, as a FIND employee put it, a firm could “ramp up manufacturing quickly enough for it to be impactful in this outbreak”—was deemed as important as the quality of the tests themselves. These considerations meant that large manufacturers with established linkages with the international organizations involved were granted priority, a bias heavily criticized in our interviews with academic units and diagnostic start-ups. Rather than a straightforward case of industry co-option of regulatory processes (Abraham and Davis 2009; Abraham and Ballinger 2012), for those involved in the international response, prioritizing “familiar entities,” spoke to the necessary trade-offs between quality and rapid distribution demanded by an emergency response. “If you are choosing between a small company and a large company,” one member of the prequalification team explained, “then you might trust the bigger one more because they are
already providing lots of tests—even if there are some performance limitations.”

Ultimately, companies with well-articulated production lines and whose R&D practices were legible to regulatory agencies presented a safer bet than tests produced by smaller entities whose quality management systems, let alone capacity to produce at scale, and were rudimentary at best. The relative imprecision of the EUAL’s benchmark of quality—a threshold that encompassed both diagnostic performance and plausible deployment—made it possible to balance the innovative potential of any singular test with the consistency and speed of supply that would render diagnosis doable in an outbreak scenario.

**Standardizing Protocols**

The efficacy of a diagnostic tool hinges on the particular protocol through which it is operationalized, and the clinical or epidemiological environment in which it is deployed. For example, the performance of a diagnostic can be affected by the prevalence of disease in the population being tested, by the reference assay against which it is evaluated, or by the type of specimen or provenance of the reagents used. Conventional metrics used to evaluate more standardized medical products, like vaccines and therapeutics, fail to capture the significance of that operational—and epidemiological—context to their public health value (Hogarth et al. 2022; Sturdy 2020). For this reason, regulatory approvals for diagnostic devices do not confer approval on a single product so much as they lay out a set of conditions for a particular device’s manufacture and use. As WHO officials were about to discover, the especially “fluid” nature of diagnostic technology (Law and Mol 2001), prone to shape-shifting and characterized by unclear object boundaries, would pose distinct regulatory challenges in an emergency context.

In November 2014, the WHO prequalification team in collaboration with the Bernard Nocht Institute, which had set up a field laboratory in Sierra Leone, ran a field assessment of the first rapid test submitted to the EUAL, the ReEBOV Antigen Rapid Test Kit manufactured by Corgenix, in collaboration with researchers from Tulane University in the United States. Upon arrival at the Bernard Nocht Institute field laboratory in Sierra Leone, however, the members of the WHO prequalification team were shocked to discover the laboratory was a repurposed classroom and that schoolbooks and desks still sat in the corner, summarily covered with bedsheets. Just establishing a protocol for running a performance evaluation under such
conditions, one team member explained to us, required improvisation, flexibility, and the adaptation of study protocols designed for a highly contained and bio-secure laboratory.

One major hurdle was obtaining blood samples on which to run the assays. Those samples, for the most part, sat in the fridges and freezers of foreign-run laboratories temporarily established to assist with the epidemic response (e.g. by Public Health England [PHE] or the US CDC), but ownership of the biological materials was ambiguous. At a minimum, running a performance evaluation study entailed applying for ethical clearance for the use of valuable blood samples from the Sierra Leone government, as well as cooperation from the laboratory where the samples were held. One, if not the main, reason why Corgenix and Tulane had been able to develop their rapid test so quickly was that they had collaborated with the Viral Hemorrhagic Fever Consortium (VHFC), located at the Kenema Government Hospital (KGH) in Sierra Leone, which had already developed a RDT for Lassa fever at the time of the Ebola outbreak (Bevan, Street, and Kelly 2018). Early in the epidemic, KEMRI operated as the only testing facility for the whole country, providing the fever consortium with an invaluable ready supply of blood samples.

As case numbers began to drop, the issue of sample access was compounded and international research teams were less inclined to share the small volumes of blood they retained. Toward the end of the outbreak, FIND ran a validation study for all four lateral flow assays submitted to the EUAL in a PHE laboratory in Sierra Leone. The study required high volumes of blood because each sample would need to be tested on five assays (the four lateral flow tests plus a reference test). Unable to source all the biological material they needed from the PHE laboratory, teams scrambled to procure frozen samples from several different laboratories across the region. However, because those samples were collected in varying amounts and their storage history was unknown, testing protocols had to be abandoned for an approach that was distinctly more ad-hoc. “Samples with a similar result from their initial PCR were pooled together, running tests on mixes of those that shared an initial PCR result,” one researcher explained. “There was just not enough volume in the samples to run them all on their own.”

As the operational costs and logistics of running evaluation studies in the midst of an outbreak became evermore apparent, the EUAL team came to rely on evidence generated by academic researchers, who tended to be associated with other international laboratories in the country and ran their own validation studies on promising tests (see, e.g., Pollock and Wonderly
2017). These researchers often fared no better: in a presentation at the McGill Global Health Diagnostics Summer School in 2018, Maya Jana Broadhurst, a microbiologist based in the United States who was involved in an independent evaluation of the Corgenix test, described her struggles:

In a very stressed and time-compressed setting, in all your PPE, when really every action must be scrutinized for safety and operability. And the challenge of “nothing leaves the red zone!” How do we collect the data and get it out of the patient-care area and get that data effectively? There was no precedent set for this kind of study.

The variability between validation protocols only generated further uncertainty on the characteristics and qualities of individual tests, raising questions as to whether those protocols could be adhered to once listed tests reached the field. Depending on what validation studies had been run, individual listings were accompanied by advice on appropriate specimen types for use (e.g., whole blood, plasma, saliva, urine), but situations could emerge where, because of available equipment, time, or patient death, those specimens were unavailable. “What does someone do in that situation?” asked one medical researcher. “Could you use it? Do you need to validate it? That is what the Food and Drug Administration would say, but what about in an emergency situation?”

**Benchmarking Performance**

By far the greatest obstacle to validation, however, was determining the reference test against which evaluations could be measured. The sensitivity and specificity of an RDTs is an artifact of its reference test. Using the Altona RealStar Filovirus Screen RT-PCR Kit version 1.0, for example, the early WHO study found the Corgenix test to have 93.8 percent sensitivity, a finding that was replicated in the FIND coordinated validation study (which showed a sensitivity of 93.18 percent). But when run against the Trombley assay, used in-house in the PHE laboratory where FIND carried out its study, it was only found to have 85.26 percent sensitivity (Wonderley et al. 2019, 7).

The viral load profiles of the samples used in the study can also generate evaluative discrepancies. This was one reason why the WHO argued for caution in response to an independent evaluation led by Broadhurst, which found the Corgenix test to have 100 percent sensitivity, higher than the RealStar reference test (Broadhurst et al. 2015; Urassa, Meurant, and Wood 2012).
Slight modifications in how the reference test was conducted could also affect results. The RealStar manufacturer Altona, for example, suggested that the reason why the Corgenix test had performed so well against their RealStar assay in Broadhurst’s study was because the team had not used the recommended DNA amplification method (Olschalger and Heb 2015). In the absence of harmonized protocols, performance evaluations were, as another medical researcher put it in an interview, often guilty of “comparing apples to oranges.”

According to the WHO’s own verification process, the Corgenix test did not meet the minimum technical specifications laid out in its TPP for Zaïre ebolavirus rapid simple test. But, as a reference point for the EUAL, the TPP was aspirational rather than prescriptive in nature. Each test also had to be considered in its own right, its unique uncertainties and weaknesses pitched against current epidemiological data, the other tools available, and immediate clinical needs. As one member of the WHO prequalification team explained:

In the EUAL process, when we were evaluating the performance, we looked at the TPP but we also looked at the situation in the field and the risk assessment. It might be a case that it doesn’t meet the TPP requirements but the risks in the field mean we say, “Okay, let’s go for it.”

In February 2015, the perceived need for a novel test that could work in remote and underresourced settings overrode uncertainties about the performance of the Corgenix test, and the WHO listed it for procurement for use in places without access to more accurate PCR (WHO, 2015c). But, as the first wave of the epidemic subsided and the diagnostic landscape changed, that calculus shifted away from the use of rapid lateral flow tests whose precision and accuracy was ill-defined, to improving access to highly accurate, automated RT-PCR nucleic acid tests. The flexibility of the EUAL mechanism, and its ambiguous status as a highly provisional guide to acceptability, allowed thresholds of tolerability, and attendant trade-offs between accuracy and expediency, to fluctuate in close alignment with the epidemic curve.

**Probable Cases**

The EUAL’s underpinning rationale was that, in the context of an emergency, urgent need outweighed the risks associated with the uncertain quality of novel technologies. But what kinds of diagnostic tests were
needed and what, precisely, they were needed for constantly changed as the epidemic unfolded. Moreover, it soon became clear that not all “communities” viewed those needs in the same way.

The WHO originally anticipated that highly portable and easy-to-use rapid lateral flow tests would be used to assist triage and quarantine decisions in rural health facilities while staff awaited the results of more accurate nucleic acid tests carried out at centralized laboratories. But by the middle of 2015, with the prevalence of Ebola dropping, there was concern that RDTs would no longer be fit for purpose. As the prevalence of disease in a population decreases, tests with lower specificity will generate a greater number of false positives. The problem with an increasing number of false positives is not only that RDTs would no longer be effective tools for determining quarantine decisions; in a context where public health authority could not be presumed, and where public opposition had caused the international response to falter, the WHO noted that there is also “a risk that use of [RDTs with low specificity] in this epidemiological context will undermine trust in the testing procedures and in the broader public health response” (WHO 2015d).

Further factors pushed the risk–benefit calculus away from lateral flow tests. Biosafety concerns around how to safely deploy, use, and dispose of PPE in primary care settings meant that, as one WHO representative explained, RDTs “could not be used outside the laboratory.” The diagnostic landscape was also changing. By mid-2015, a growing array of novel automated RT-PCR platforms were serving a nationwide network of laboratories and Ebola Treatment Units with highly accurate tests, at a time when falling case numbers had removed the need for prescreening. In late 2015, the WHO advised against the deployment of rapid lateral flow antigen tests (WHO 2015d). By the time the FIND team traveled to Sierra Leone to run a performance evaluation of the four lateral flow RDTs submitted to the EUAL, the purpose was not to establish whether any of the rapid lateral flow tests could fulfill current needs but rather, as one researcher explained, to determine “possible use cases” for the tests in future outbreaks.

But even as the WHO and FIND readjusted their diagnostic aspirations, those working in Sierra Leone reasserted the continuing value of RDTs for the response. The response team in Kono District, for instance, had used the Corgenix test in the early days of the outbreak to screen suspected patients, enabling health workers to rapidly isolate those who were more likely to be infected. When government officials in Freetown learned that Corgenix was being used for this purpose, they ordered an immediate end to the practice,
in line with WHO guidance. For the response team in Kono, the disconnect between those top-down guidelines and diagnostic needs on the ground was profound:

For us, it was like a setback. We were in the field, we were quite confident that it was helping us a lot to differentiate highly probably cases from the less probable cases. Not that we were announcing it that it was positive or not, but it was helping us in the field to actually differentiate who is likely to be positive and who is likely not to be positive. So that was helping us to do that. We really frowned at the fact that the WHO and others said we should stop using it. In fact, at one point I said, “No, we are going to use it, as actually in the field it’s helping to protect my team,” because I led a team to Kono. [I said] “It’s good to protect our team so we are going to use it.” We defied whatever orders came from Freetown as long as we are not announcing the result. [My response was] “It was not meant for you; it is meant to protect us in the field.” So we are using it.

The identification of a “probable case” articulated a novel use case for Corgenix—one that made its good-but-not-perfect performance more tolerable. That clinical use, however, was not one that the WHO could ultimately support. The WHO’s decision to discourage the use of RDTs outside the laboratory was guided by a set of epidemiological considerations increasingly at odds with the needs identified by frontline workers, described in an interview for this study:

There was a lot of disagreement about how to implement something that was not perfect. People were working in the field with patients and they wanted to use it, even if it wasn’t perfect. But WHO and FIND didn’t want to use a test that had a sensitivity/specificity deficit. They didn’t want a false positive to throw everything into disarray. Which seems fair enough from one perspective. But it also left clinical units without testing facilities. They still had to transport the samples and wait for the results. And this was never really resolved.

In the particular circumstances of the West African Ebola outbreak—in an under-resourced healthcare system overwhelmed (and arguably undermined) by the organizational complexity of the international response—challenges in assessing the scientific credibility of manufacturers’ claims were exacerbated by countervailing imaginaries of diagnostic purpose. The WHO’s sense of what constituted a reasonable use case corresponded with epidemiological assessments that increasingly diverged from the day-to-
day clinical contexts of use. For those involved in the response, diagnostic value had to be adjudicated on the ground, demanding an experiential and highly situated expertise that did not align neatly with the evidentiary processes the WHO was trying to put in place. These fault lines highlighted the failure of the emergency regulatory response to adequately involve potential users—arguably the constituency with the most at stake in the deployment of novel diagnostic tools—in the determination of either diagnostic needs or thresholds of tolerability.

Quasi-regulation

“An even bigger issue for the EUAL,” said one FIND employee, “was: what authority did it convey?” With power neither to compel nor enforce rules in member states, the WHO’s regulatory weight derives from its technical capacities; its role is to provide guidance, coordination, and supply chain oversight in donor markets for medical products in low- and middle-income countries. During the Ebola outbreak, listing through the EUAL process was a prerequisite for product procurement by UN agencies and their affiliates, an imprimatur of quality to support large purchase orders from organizations such as UNICEF.

WHO (2015e, 6) guidelines for the use of in vitro diagnostics (IVD) stated that “regulation of IVDs is normally the role of the national regulatory authority,” yet the EUAL was also intended as a guide for National Regulatory Authorities (NRAs) without the capacity to carry out their own validation procedures. From the perspective of those NRAs, however, the EUAL could be experienced as a trojan horse—a means for international researchers, global health policymakers, and commercial developers to circumvent NRAs’ sovereign authority under the banner of technical guidance. While the EUAL formally deferred to NRAs in matters of in-country regulation, the international response in fact provided little actual assistance for governments in afflicted countries to strengthen their national capacity for regulation.

Government officials working in pharmacovigilance for the Pharmacy Board of Sierra Leone, for instance, complained that they received no additional resources during the outbreak to assist with the processing of numerous applications for IVDs submitted by international research groups, nongovernmental organizations, and UN agencies. This resulted in substantial workloads and major backlogs, but nonetheless Sierra Leone’s NRA was under intense pressure from international organizations to authorize the use of novel products, especially once they had been listed by the EUAL.
While international organizations like the WHO and FIND formally bowed to the legal authority of NRAs, they sometimes claimed the moral authority to act in exceptional circumstances. With long waiting times for approval from NRAs in West Africa, one person involved in the diagnostic response acknowledged that “some tests were being brought in and used immediately just because of the sense of urgency of the outbreak.” Meanwhile, international laboratories often procured their equipment and materials privately, bypassing both the EUAL and national regulators. Sierra Leone and WHO officials also expressed frustration that international laboratories that were ostensibly established to deliver urgent assistance were in fact trafficking blood samples to university and private laboratories in Europe and North America for the purpose of IVD product development—suspicions that were ultimately borne out (Freudenthal 2019).

For officials in Sierra Leone, the evasion of national regulations by international organizations and research institutions alike was experienced as endemic, symptomatic of those organizations’ lack of respect for national sovereignty during the emergency. As for Benton’s (2015) analysis of the foreign-sponsored programs for HIV/AIDS treatment, “good governance” for the Ebola response was measured in the speed and incisiveness of international interventions, which often operated in exception to, rather through integration with, the public health system. Toward the end of the outbreak and in its immediate aftermath, as media reporting, stakeholder feedback and internal performance reviews forced the WHO to contend with these critiques, the need to reexamine the qualities of the relationships through which emergency response is achieved became starkly apparent. As the WHO (2016, 11) put it in “An R&D Blueprint for Action to Prevent Epidemics” published soon after the end of the outbreak: “The [Ebola] research response suffered from insufficient local technical capacity and deficient understanding by international partners of the fundamental needs and culture of the affected communities.”

Conclusion

The Ebola outbreak was, the WHO later reflected, a crisis for which the international community was woefully underprepared, with “no vaccines, no treatments, few diagnostics, and insufficient medical teams and trained responders” (WHO 2016, 6). In the face of that profound medical scarcity, the EUAL provided a novel international regulatory platform to incentivize the development of lifesaving products and accelerate their
deployment—a means of extending a provisional avenue of moral author-ization, if not legal authority.

Measured against the goal of accelerating access to diagnostics within the time frame of the outbreak, the EUAL had mixed results. Many of the nucleic testing platforms it listed contributed substantially to the response, providing rapid results in Ebola Treatment Units around the country and enabling clinicians to make decisions about quarantine and treatment with confidence. Some of those platforms remained in the country after the outbreak ended and were repurposed for testing for other diseases. However, their contribution to laboratory strengthening has been undermined by the unavailability and/or unaffordability of test cartridges, while refresher training in PCR methods for laboratory staff is scarce (Ansumanna et al. 2020; Vernooij 2021). Even more ignominious is the fate of RDTs. The failure to establish a consensus around a viable use case and dependable market has meant that none of the devices initially listed continued to be commercially available for subsequent outbreaks across the African region (Cnops et al. 2019).

Yet those failures should not fully impugn the potential of the EUAL for global health innovation. Instead, we suggest that the EUAL is best understood as providing a liminal space of regulation, an “in-between” process through which the norms and standards of a new regime of emergency regulation were worked out in real time. For the duration of the EUAL for Ebola diagnostics, the actors involved in its implementation found themselves navigating multiple fault lines: between clinical accuracy and reliability versus speed and scale of manufacturing; between market-based logics and humanitarian needs; between national and international regulatory regimes; and between the desire to meet the diagnostic requirements of the current outbreak versus the need for readiness of effective diagnostics for future epidemics. What could justifiably be done in the name of an emergency was continually redrawn and circumscribed by that volatile moral calculus over the course of the response. In this context, the seemingly scientific parameters of risk–benefit calculations (such as establishing the efficacy and safety of a novel diagnostic test) overflowed into social, legal, ethical, and political terrain. Epidemiological unknowns became entangled with ethical unknowns: facts were unsettled but values were also in dispute (Funtowicz and Rayetz 1993, 1882). The collision of science, diplomacy, and ethics in the EUAL process revealed the multiple, overlapping interests and risks associated with emergency regulation (Nuffield Council on Bioethics 2020) but also made “the network of problems it raised both visible and debatable” (Callon, Lascoumes, and Barthe 2011, 31).
In science studies, there is a long tradition of research into the democratic potential of public reckoning with the values of science that may diverge from expert views, such as biotechnological controversies, “hybrid forums” (Callon, Lascoumes, and Barthe 2011), “reflexive institutions” (Fischer 2003), and “postnormal science” (Funtowicz and Rayetz 1993). During these critical events, regulatory science undergoes a transformation, expanded to include multiple stakeholders and shifting from “formalized deduction to an interactive dialogue” (Funtowicz and Rayetz 1993). When both systemic uncertainties and the stakes of decisions are high, conflicting values become impossible to ignore, opening the way for “the legitimacy of a plurality of perspectives and ways of knowing” (Funtowicz and Rayetz 1993, 1885).

To suggest that the challenges encountered in the implementation of the EUAL prompted a wholehearted move toward pluralism and “postnormal” decision-making in the WHO’s emergency regulation procedures would perhaps overstate the flexibility of a procedure designed primarily to incentivize industry to develop novel products. Nonetheless, the WHO’s approach to the EUAL proved to be highly reflexive, its shortcomings and failures prompting processes of review, consultation, deliberation, and change. For instance, a 2017 WHO consultation on regulatory preparedness grappled with its limitations as a regulatory body and recommended that the EUAL be renamed Emergency Use Listing (EUL) to “clarify the relationship of the EUAL process to activities at the NRA [National Regulatory Authority] level and explain how the EUAL may be used by NRAs and procurers in their decision making” (WHO 2017, 6; see also Perkins et al. 2017). The revised procedure emphasized the role of NRAs in approving research protocols and signing off on the transport and export of samples. It also placed the onus on the WHO to share relevant presubmission data with national governments and to support and coordinate in-country ethical reviews (Laurie et al. 2018; Sethi 2018). Perhaps most strikingly, the review recognized the ongoing ethical challenges surrounding access to biological materials for product development and validation, pointing to the lack of clarity around ownership, and the need for mechanisms that can both protect the interests of epidemic-afflicted countries while ensuring that materials are shared quickly in public health emergencies.

In 2016, the WHO launched the R&D Blueprint and in so doing took on a new leadership role in accelerating and coordinating emergency research and product development (Moon et al. 2015; Mehand et al. 2018). The successive reformulation of protocols for emergency R&D as part of that
agenda suggests that those norms are fluid and evolving. The EUAL, now EUL, continues to make emergencies doable in particular ways, even if only in showing the limitations of certain courses of international action and in providing a forum for a more inclusive ethical debate. But for the R&D blueprint to make good on that democratic promise, pregiven assumptions about needs and goals must not overdetermine whose voices are heard and whose expertise is sought. As Melissa Leach and colleagues have forcefully argued, acknowledging and managing uncertainty in public health emergencies requires “project and programme styles that emphasise learning, iterative adaptation, flexible action and equitable relationships amongst diverse actors [that] need to move from the margins to centre-stage” (Leach et al. 2021, 6).

A more tolerable and supple emergency science would not be defined within the narrow purview of which products best suit preestablished technical needs but would instead start with a reflexive and inclusive discussion about how diagnostic needs might be defined; who and what diagnostics are for; and how those needs might fluctuate and change. Put differently, rather than create legal silos of regulatory attention around artificial and fixed diagnostic objects, the EUAL attends to the “spaces in-between” the current regulatory environment for health R&D (Laurie 2017). That liminality, we show, is a crucial legacy of the Ebola outbreak for the WHO, an organization whose limited power to compel or enforce guidance puts it in the unique position to expand global health governance to a far wider range of actors, not only product developers or manufacturers but research, national, and local officials, laboratory experts, and health workers on the front line of emergency response.

Acknowledgments
We are grateful to the wider team who contributed to the research on which this article is based. Eva Vernooij undertook one of the interviews that we draw on in this article and provided feedback on early drafts. Shona Lee undertook research at an international workshop that we draw on, and Imogen Bevan carried out a review of gray literature that informed our analysis and provided feedback on early drafts. We are grateful to our many interlocutors, who gave us their time in the midst of exceptionally busy schedules.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program under grant agreement No 715450. AHK acknowledges the support provided by the Princeton Institute for Advanced Study Social Science Fellowship, “Platform Theme”, 2023-2024.

ORCID iDs
Alice Street https://orcid.org/0000-0001-7874-0234
Ann H. Kelly https://orcid.org/0000-0003-4658-2133

Notes
1. Drawing from Victor Turner’s (1969) classic anthropological studies of coming-of-age rituals, Laurie (2017) offers “the liminal spaces of health regulation” as a framework to engage with the processual dynamics of health research in a way that can better accommodate the diversity of activities, stakeholders, and the experiences involved (Laurie 2017, 8; see also Sethi 2018; Ganguli-Mitra and colleagues 2017).

2. This article draws together different strands of research undertaken in the context of the European Research Council–funded Investigating the Design and Use of Diagnostic Devices in Global Health (DiaDev) project. The DiaDev team has undertaken more than eighty interviews with product developers, public health experts, emergency responders, local patients, and medical professionals in the United States, Europe, and Sierra Leone to understand the role novel diagnostic devices play in responding to public health emergencies. The insights from those interviews have been contextualized by reviews of gray and scholarly literature regarding diagnostic development and situated within an extensive document analysis of gray literature generated by the WHO, ministries of health, and other global health organizations as well as by attendance at conferences and meetings in the global health diagnostics field. All research data drawn on in this article were collected with ethical approval from the School of Social and Political Science Research and Research Ethics Integrity Committee at the University of Edinburgh and the Sierra Leone Ethics and Scientific Review Committee. We followed best practices in anthropology regarding informed consent.

3. Increasing access to and stimulating the markets for medical products are priority acts in global health security: when a public health problem is deemed an international threat, global health policy and investment trends privilege interventions that can be stockpiled (Roemer-Mahler and Elbe 2016). In the case of the West Africa Ebola outbreak, however, the commercial development and deployment
of novel medical tools were defined by the humanitarian needs of the affected population rather than by national biosecurity strategy.

4. A precedent for the relaxation of regulatory requirements during public health emergencies was set by the Federal Drugs Authority (FDA) in the United States in the early 2000s, when it introduced the Emergency Use Authorization (EUA) procedure for medical products in response to emerging bioterrorism threats. The procedure can only be activated following a declaration by the Secretary of Health and Human Services that emergency circumstances exist that warrant the introduction of the EUA. This enables the use of unapproved medical products that “may be effective” and for which there is no available alternative for the duration of that emergency (FDA 2017).

References


Perkins, Mark D., Christopher Dye, Manica Balasegaram, Christian Bréchet, Jean-Vivien Mombouli, John-Arne Røttingen, Marcel Tanner, and Catharina C.


Author Biographies

Alice Street is a Professor of Anthropology and Health at the University of Edinburgh. She has published widely on the anthropology of health system strengthening, global health interventions, and medical innovation for public health emergencies and is the author of Biomedicine in an Unstable Place: Infrastructure and Personhood in a Papua New Guinea Hospital, published by Duke University Press. Professor Street is Principal Investigator of DiaDev, a six-year international research program to investigate the design and the use of diagnostic devices in global health, funded by the European Research Council, which examined the changing role of diagnostics in the context of public health emergencies, health system strengthening, and disease elimination.

Ann H. Kelly is a Professor of Anthropology and Global Health at Kings College London. Her ethnographic work focuses on the socio-material conditions that structure the production of global health knowledge and the local ecologies of labor that circumscribe its circulation and use. She explores these themes in a forthcoming book with Duke University Press on pasts and futures of mosquito control, with Javier Lezaun. She serves as a member of the WHO Strategic Advisory Group of Experts (Sage) for Ebola Vaccines and Vaccination and as a member of the WHO Strategic Advisory Group of Experts (Sage) for Ebola Vaccines and Vaccination.