Open Forum Infectious Diseases

MAJOR ARTICLE

HIV Testing After a First Positive Rapid Diagnostic Test: A Role for Nucleic Acid Testing?

Anne M. Neilan,[1,2] Jennifer E. Cohn,[3,4] Jean-Francois Lemaire,[3] Emma Sacks,[3,5] Rebecca Alban,[3] Kenneth A. Freedberg,[1,6] Rochelle P. Walensky,[1,6] and Andrea L. Ciaranello[1]

[1]Division of Infectious Diseases and Medical Practice Evaluation Center, Massachusetts General Hospital, and Harvard Medical School, Boston, Massachusetts; [2]Department of Pediatrics, Massachusetts General Hospital Boston, Massachusetts; [3]Elizabeth Glaser Pediatric AIDS Foundation, Geneva, Switzerland; [4]Division of Infectious Diseases, University of Pennsylvania, Philadelphia, Pennsylvania; [5]Johns Hopkins School of Public Health, Baltimore, Maryland; [6]Division of General Internal Medicine, Massachusetts General Hospital, Boston, Massachusetts

We developed an open-access, Excel-based model simulating currently recommended and alternative algorithms for adult HIV testing as a preliminary investigation of trade-offs between accuracy and costs. Despite higher costs, simpler HIV testing algorithms incorporating point of care nucleic acid testing may improve outcomes and thus merit additional research and field testing.

Keywords. accuracy; costs; human immunodeficiency virus; nucleic acid tests; rapid diagnostic tests.

Toward the goal of diagnosing 90% of people with HIV by 2030, the World Health Organization (WHO) outlined an adult HIV testing algorithm based on rapid diagnostic tests (RDTs). However, uptake of the current WHO algorithm is problematic—less than 20% of country-level HIV testing protocols follow WHO recommendations [1, 2], likely due to the algorithm's complexity and number of required assays [3]. For example, up to 8 RDTs may be needed (Supplementary Figure 1: discordant RDT results, low prevalence settings).

Algorithm accuracy is also concerning: program audits report many false-positive (range, 0%–10%) [4–8] and false-negative results (7%) [4, 9]. Incorrect results have a substantial impact on individuals and lead to errors in estimation of program requirements and costs. Point of care (POC) nucleic acid testing (NAT) may facilitate antiretroviral therapy (ART) initiation by permitting same-day HIV diagnosis but may have higher per-assay costs. As a preliminary investigation of accuracy and costs, we modeled simpler alternative algorithms incorporating POC-NAT, compared with the current RDT-based WHO HIV testing algorithm [1].

METHODS

Using an Excel-based model, we simulated 3 currently recommended and novel alternative algorithms for adult HIV diagnostic testing (Supplementary Figure 1): current RDT-based WHO guidelines; POC-NAT to resolve discordant RDT results (NAT-resolve); and POC-NAT to confirm a first positive RDT result (NAT-confirm). In the online Excel tool, model-users may specify HIV prevalence; assay sensitivity, specificity, and costs; and the lifetime per-person cost of HIV care and ART ("care/ART cost") to reflect a range of settings. We assumed conditional independence across tests (ie, that there is no reason 1 false-positive will result in another) [3, 10]. In practice, patients who receive a final "inconclusive result" should repeat the entire algorithm; in the model in these cases, we applied an average per-person algorithm assay cost again and assigned results for repeat testing based on true infection status. We assumed that the NAT-resolve and NAT-confirm algorithms would not include any pre-ART retesting, based on low uptake of retesting recommendations in practice [3].

In the base case, we reflected high-burden, low-income settings: 15% HIV prevalence and lifetime care/ART costs of $4000, estimated from model-based analyses (Supplementary Table 1) [11]. We assumed that after false-positive diagnosis, patients incurred lifetime care/ART costs equivalent to after true-positive diagnosis; after false-negative diagnosis—with later risk of costly opportunistic infections or transmission of HIV without ART—patients incurred HIV care costs equal to 75% of lifetime care/ART costs [12]. Assay sensitivity/specificity/cost were derived from averaging WHO Prequalification Reports and from UNICEF (RDT: 99.8%/99.6%/$1.47; POC-NAT: 95.0%/100%/$27.92) [13, 14]. Outcomes included accuracy ([true-positive + true-negative results]/all results), number and costs of assays required, and care/ART costs. We excluded all clinical outcomes, including morbidity, mortality, onward transmissions and related costs, associated with true- and false-positive and -negative results.

In scenario analyses, we simulated combinations of prevalence and cost to reflect diverse settings: high-prevalence (15%) and low-prevalence (3%), with middle-income ($20 000) and low-income ($4000) lifetime care/ART cost. We also simulated a field-based setting, with lower test sensitivity/specificity, as

Received 6 February 2018; editorial decision 11 February 2018; accepted 12 July 2018.

Correspondence: A. M. Neilan, MD, MPH, 100 Cambridge Street, Floor 16, Boston, MA 02114 (aneilan@partners.org).

Open Forum Infectious Diseases®

DOI: 10.1093/ofid/ofy170

might occur when implemented in the field (RDT: 89.0%/82.9%; POC-NAT: 94.0%/99.4%) [5, 15].

In sensitivity analyses, we examined a range of parameters, including the cost of false-negative diagnoses (up to 7-fold true-positive costs) and specific assays (up to 30-fold base case). Additional model inputs, assumptions, and testing algorithms (including distinct WHO algorithms for low-prevalence [<5%] or high-prevalence [≥5%] settings) are available in the Supplementary Appendix.

RESULTS

Overall Accuracy

In the base case, NAT-resolve had similar accuracy as the WHO algorithm (99.97%); NAT-confirm was least accurate (99.23%) (Table 1).

False-Negative Results

NAT-confirm led to the most false-negative diagnoses (774/100 000) due to low POC-NAT assay sensitivity (94.00%). When POC-NAT was instead used to resolve discordance between 2 RDTs (NAT-resolve), the effect of the NAT's lower sensitivity was mitigated by the greater sensitivity of the second RDT (99.83%): 27/100 000 were false-negative results, similar to 26/100 000 for the WHO algorithm (Table 1).

False-Positive Results

NAT-confirm led to 0% false-positive results. Among the other strategies, false-positive results were similar (1–3/100 000 false-positive results) (Table 1).

Assay Utilization and Costs

On average, the WHO algorithm required more assays/person compared with all other strategies (1.5 vs 1.2) (Table 1). Costs attributable to the assays alone were greatest for NAT-confirm ($5.75/person tested), whereas assay cost for NAT-resolve ($1.80/person tested) was closer to, but still lower than, the WHO algorithm ($2.15/person tested). The total algorithm cost (including lifetime care/ART) for NAT-resolve was similar to

Table 1. Base Case and Setting Analyses: Modeled Outcomes of 3 HIV Testing Strategies

HIV Testing Strategy[1]	Assay Performance			Assays/Person Tested[4]	Algorithm Assay Cost/Person Tested, $[4]	Total Cost/Person, $[4]	Strategy Cost Compared With WHO Algorithm, $[5]
	Accuracy, %[2]	False-Negative per 100 000[3]	False-Positive per 100 000[3]				
Base case: high-prevalence, low-income (15% HIV prevalence, $4000 lifetime HIV care and ART costs)							
WHO algorithm	99.97	26	2	1.5	2.15	602	Reference
NAT-resolve	99.97	27	1	1.2	1.80	602	0
NAT-confirm	99.23	774	0	1.2	5.75	598	–4
High-prevalence, middle-income (15% HIV prevalence, $20 000 lifetime HIV care and ART costs)							
WHO algorithm	99.97	26	2	1.5	2.15	3001	Reference
NAT-resolve	99.97	27	1	1.2	1.80	3001	0
NAT-confirm	99.23	774	0	1.2	5.75	2967	–34
Low-prevalence, low-income (3% HIV prevalence, $4000 lifetime HIV care and ART costs)							
WHO algorithm	99.99	5	3	1.2	1.71	122	Reference
NAT-resolve	99.99	5	2	1.0	1.63	122	0
NAT-confirm	99.85	155	0	1.0	2.41	121	–1
Low-prevalence, middle-income (3% HIV prevalence, $20 000 lifetime HIV care and ART costs)							
WHO algorithm	99.99	5	3	1.2	1.71	602	Reference
NAT-resolve	99.99	5	2	1.0	1.63	602	0
NAT-confirm	99.85	155	0	1.0	2.41	595	–7
Field-based setting: lower testing sensitivity/specificity (15% HIV prevalence, $4000 lifetime HIV care and ART costs)							
WHO algorithm	93.10	1982	4914	2.0	2.96	698	Reference
NAT-resolve	95.72	1723	2558	1.4	5.65	691	–7
NAT-confirm	97.60	2318	87	1.3	9.26	590	–108

Relative to the WHO algorithm, green indicates a more favorable result; orange indicates a less favorable result; blue indicates a neutral result.

Abbreviations: ART, antiretroviral therapy; NAT, nucleic acid test; WHO, World Health Organization.

[a]Base case assay sensitivity/specificity/cost were derived from averaging WHO Prequalification Reports and the UNICEF catalog (rapid diagnostic test [RDT]: 99.8%/99.6%/$1.47; point of care [POC]–NAT: 95.0%/100%/$27.92) [13, 17]. Field-based sensitivity/specificity/cost were derived from published reports (RDT: 89.0%/82.9%; POC-NAT: 94.0%/99.4%) [5, 15]. Relative to the other strategies, differences in base case per-person total costs did not change when differing RDT sensitivity/specificity/cost were substituted for each step in an algorithm: when substituted for A1: Determine 100%/98.93%/$1.43; A2: SD-Bioline 100%/99.9%/$1.07; A3: Uni-Gold 100%/98.93%/$1.93 [13]. (Results may be replicated in the available online Excel tool, not shown here.)

[b]Accuracy is [(true-positive + true-negative)/ (true-positive + false-positive + false-negative + true-negative)]/person tested.

[c]Proportion of all test results that are false-negative (third column) or false-positive (fourth column) per 100 000 people tested. This may also be expressed as a percentage.

[d]Total number of assays (fifth column) required or costs attributable to assays for an algorithm (sixth column) or lifetime ART and HIV care costs (seventh column) per person tested (including HIV-infected and HIV-uninfected people tested). Supplementary Figure 2, A–B present sensitivity analyses varying the cost for false-negative diagnoses in the total cost per-person tested.

[e]High (15%) HIV prevalence settings are compared with the WHO high-prevalence algorithm. Low (3%) HIV prevalence settings are compared with the WHO low-prevalence algorithm. NAT-confirm was costlier than the WHO algorithm when POC-NAT cost was >$55; NAT-resolve was costlier than the WHO algorithm when POC-NAT cost was >$270. (See Supplementary Table 2. Results may be replicated in the available online Excel tool, not shown here.)

the WHO algorithm, whereas NAT-confirm cost less than the WHO algorithm by $4/person tested.

Sensitivity Analyses

Holding all other parameters equal to the base case, NAT-confirm became more costly than the WHO algorithm when the POC-NAT cost was >$55 (Supplementary Table 2) or when lifetime HIV care cost after false-negative diagnosis was equal to or exceeded lifetime care/ART cost after true-positive diagnosis (Supplementary Figure 2A).

Alternative Prevalence and Income Settings

In low-prevalence settings, NAT-confirm and NAT-resolve still required fewer assays/person than the WHO algorithm: 1.0 vs 1.2 assays/person tested (Table 1). In both high- and low-income settings, NAT-resolve had accuracy and total algorithm cost similar to the WHO algorithm, and NAT-confirm was less accurate and less expensive. For all algorithms, overall costs varied widely from lifetime ART/care cost ($121–$3001).

Field-Based Settings

Substituting lower field-based sensitivities/specificities for all assays, NAT-confirm became the most accurate algorithm (97.60% vs 93.10%–95.72%), in contrast to the base case. It was also the least expensive ($590/person vs $691–$698/person) (Table 1). NAT-confirm's greater overall accuracy included fewer false-positive results (87 vs 4938/100 000) but more false-negative results (2318 vs 1982/100 000) than the WHO algorithm. NAT-confirm remained less expensive than the WHO algorithm, except when the POC-NAT cost was >15-fold higher than the base case ($275) (Supplementary Table 2) or when lifetime HIV care cost after false-negative diagnosis was >7-fold higher than lifetime ART/care cost after true-positive diagnosis (>$30 000) (Supplementary Figure 2B). NAT-resolve became costlier than the WHO algorithm when the POC-NAT cost was >3-fold higher than the base case ($85) (Supplementary Table 2); NAT-resolve was less expensive than the WHO algorithm at all examined costs of false-negative diagnosis (Supplementary Figure 2B) in the field-based setting.

DISCUSSION

We find that compared with the WHO algorithm, POC-NAT-based algorithms (NAT-confirm and NAT-resolve) may require fewer assays and have similar or lower costs. Using WHO Prequalification Report sensitivity/specificity, NAT-resolve had equal accuracy and lower lifetime costs than current WHO algorithms; using field data, however, NAT-confirm was more accurate and less expensive than WHO algorithms.

Despite concerns about assay cost, this was not the most influential parameter: POC-NAT assay cost had to increase 2–15-fold for NAT-based algorithms to become costlier than the WHO algorithm. Importantly, the cost savings with both NAT-based strategies reflect 2 opposing effects: permitting false-negative results (less expensive because treatment is not immediately provided, but risking higher morbidity and mortality) and averting false-positive results (avoiding unnecessary care/ART costs for people without HIV, potentially for life). NAT-confirm is most cost-saving in the base case, but its low cost is driven primarily by failure to identify HIV-infected people (false-negative results)—clearly a harmful outcome of this approach. This focused model excludes key long-term clinical, transmission, and cost outcomes of incorrect diagnoses, and thus is intended to inform further research rather than provide definitive guidance.

Our findings highlight 2 important gaps in the currently available data: First, reproducible estimates of assay sensitivity and specificity in the field are needed. The base case and field scenario results differ due to differences in reported sensitivity/specificity, which are the most influential determinants of model results. Notably, the field-based modeled results for the WHO algorithm closely match published proportions of false-positive (5%) and false-negative (2%) results (Table 1) [4–8]. When applied at scale, decimal differences in sensitivity/specificity have substantial implications: in the high-prevalence base case, for every 10 000 people tested, a 1% decrement in sensitivity causes 14 additional false-negative results, and a 1% decrement in specificity causes 3 additional false-positive results. WHO guidelines assume assay sensitivity/specificity comparable to values in the WHO Prequalification Report—at least 99/98% [1], whereas published ranges reach far lower (89.0/82.86%) (Supplementary Table 1) [5, 8, 13, 15, 16]. Second, clinical studies and detailed simulation modeling of incorrect diagnoses are needed to more fully capture resulting clinical and economic outcomes. In the absence of these data, we excluded important clinical outcomes and their associated costs, but this information will be critical to inform algorithm selection.

In addition to excluding clinical outcomes, this model makes key simplifying assumptions: First, for NAT-based algorithms, we exclude some potential costs (eg, training or equipment costs for implementing and maintaining a new testing algorithm or pre-ART retesting) or additional benefits (eg, increasing access to POC-NAT for early infant diagnosis or enabling same-day confirmed HIV diagnosis and linkage to care and ART). Second, the model assumes perfect algorithm fidelity; as a result, the benefits of complex WHO algorithms may be overestimated. Finally, assays may not be conditionally independent as modeled [3], and some countries may use assays with widely differing sensitivity and specificity for sequential or confirmatory testing, which may change the overall accuracy and costs (Table 1) [1, 8].

High assay costs are often cited as a rationale for not exploring NAT-based algorithms. We find that assay cost has minimal impact on overall algorithm costs relative to assay sensitivity and specificity and the costs of lifetime HIV care and ART. Detailed assessment of assay performance and cost in the field, as well as of the clinical harms and costs associated with false-positive

and false-negative results, is needed to inform future guideline development. Simpler HIV testing algorithms incorporating POC-NAT may improve outcomes and thus merit additional research and field testing. In the meantime, we demonstrate that HIV testing algorithms integrating POC-NAT should not be dismissed simply because of their perceived high per-test cost before algorithm fidelity, accuracy, and downstream costs are also critically examined.

Supplementary Data

Supplementary materials are available at *Open Forum Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Acknowledgments

Authors' contributions. All authors contributed substantively to this manuscript in the following ways: study design (A.M.N., J.E.C., A.L.C.), development of the Excel model (A.M.N., A.L.C.), data analysis (A.M.N., J.E.C., A.L.C.), interpretation of results (all authors), drafting the manuscript (A.M.N., A.L.C.), critical revision of the manuscript (all authors), and final approval of submitted version (all authors).

Financial support. This work was supported by the Harvard University Center for AIDS Research (P30AI060354 to A.M.N.), the International Maternal Pediatric AIDS Clinical Trials Network (UM1AI068632 to A.M.N.), the Eleanor and Miles Shore Scholars in Medicine Fellowship (to A.M.N.), the National Institute for Allergy, Immunology and Infectious Diseases (R01 AI058736 to K.A.F.; R37 AI093269 to R.P.W.), the Steve and Deborah Gorlin MGH Research Scholars Award (to R.P.W.), the National Institute for Child Health and Development (R01 HD079214 to A.L.C.), and the Elizabeth Glaser Pediatric AIDS Foundation (0017A to A.L.C.).

Disclaimer. The findings and conclusions in this article are those of the authors and do not necessarily represent the official position of any of the funding organizations.

Potential conflicts of interest. The authors declare that they have no competing interests. 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. Johnson C, Fonner V, Sands A, et al; World Health Organization. Consolidated guidelines on HIV testing services. Annex 7: detailed description of HIV in vitro diagnostic formats; and Annex 14: a report on the misdiagnosis of HIV status. **2015**. Available at: http://apps.who.int/iris/bitstream/10665/180231/1/WHO_HIV_2015.33_eng.pdf?ua=1. Accessed 8 December January 2018.
2. World Health Organization. World Health Organization information note: reminder to retest all newly diagnosed HIV-positive individuals in accordance with WHO recommendations. **2014**. Available at: http://www.euro.who.int/en/health-topics. Accessed 8 January 2018.
3. Klarkowski D, O'Brien DP, Shanks L, Singh KP. Causes of false-positive HIV rapid diagnostic test results. Expert Rev Anti Infect Ther **2014**; 12:49–62.
4. Kufa T, Lane T, Manyuchi A, et al. The accuracy of HIV rapid testing in integrated bio-behavioral surveys of men who have sex with men across 5 Provinces in South Africa. Medicine **2017**; 96:1–7.
5. Boadu R, Darko G, Nortey P, et al. Assessing the sensitivity and specificity of first response HIV-1-2 test kit with whole blood and serum samples: a cross-sectional study. AIDS Res Ther **2016**; 13:1–8.
6. Klarkowski DB, Wazome JM, Lokuge KM, et al. The evaluation of a rapid in situ HIV confirmation test in a programme with a high failure rate of the WHO HIV two-test diagnostic algorithm. PLoS One **2009**; 4:1–6.
7. Shanks L, Klarkowski D, O'Brien DP. False positive HIV diagnoses in resource limited settings: operational lessons learned for HIV programmes. PLoS One **2013**; 8:1–6.
8. Shanks L, Siddiqui MR, Kliescikova J, et al. Evaluation of HIV testing algorithms in Ethiopia: the role of the tie-breaker algorithm and weakly reacting test lines in contributing to a high rate of false positive HIV diagnoses. BMC Infect Dis **2015**; 15:1–10.
9. Wolpaw BJ, Mathews C, Chopra M, et al. The failure of routine rapid HIV testing: a case study of improving low sensitivity in the field. BMC Health Serv Res **2010**; 10:1–4.
10. Kosack CS, Shanks L, Beelaert G, et al. Designing HIV testing algorithms based on 2015 WHO guidelines using data from six sites in Sub-Saharan Africa. J Clin Microbiol **2017**; 55:3006–15.
11. Bendavid E, Young SD, Katzenstein DA, et al. Cost-effectiveness of HIV monitoring strategies in resource-limited settings: a southern African analysis. Arch Intern Med **2008**; 168:1910–8.
12. Tagar E, Sundaram M, Condliffe K, et al. Multi-country analysis of treatment costs for HIV/AIDS (MATCH): facility-level ART unit cost analysis in Ethiopia, Malawi, Rwanda, South Africa and Zambia. PLoS One **2014**; 9:1–11.
13. Viljoen J, Gampini S, Danaviah S, et al; World Health Organization/ANRS 1289 Kesho Bora Study Group. Dried blood spot HIV-1 RNA quantification using open real-time systems in South Africa and Burkina Faso. J Acquir Immune Defic Syndr **2010**; 55:290–8.
14. Fitzgerald N, Cross M, O'Shea S, Fox J. Diagnosing acute HIV infection at point of care: a retrospective analysis of the sensitivity and specificity of a fourth-generation point-of-care test for detection of HIV core protein p24. Sex Transm Infect **2017**; 93:100–1.
15. Kroidl I, Clowes P, Mwalongo W, et al. Low specificity of determine HIV1/2 RDT using whole blood in south west Tanzania. PLoS One **2012**; 7:e39529.
16. World Health Organization. In vitro diagnostics and laboratory technology: prequalification of in vitro devices and medical devices. Available at: http://www.who.int/diagnostics_laboratory/evaluations/en/. Accessed 8 January 2018.
17. UNICEF supply catalog: HIV test kits. Available at: https://supply.unicef.org/. Accessed 8 January 2018.