





Comparison of Detection Limits of Fourth- and Fifth-Generation Combination HIV Antigen-Antibody, p24 Antigen, and Viral Load Assays on Diverse HIV Isolates

 Mars Stone,^{a,b} John Bainbridge,^{c,d} Ana M. Sanchez,^{c,d} Sheila M. Keating,^{a,b} Andrea Pappas,^{c,d} Wes Rountree,^{c,d} Chris Todd,^{c,d} Sonia Bakkour,^{a,b}  Mark Manak,^e Sheila A. Peel,^f Robert W. Coombs,^g Eric M. Ramos,^g M. Kathleen Shriver,^h Paul Contestable,ⁱ Sangeetha Vijaysri Nair,^j David H. Wilson,^k Martin Stengelin,^l Gary Murphy,^m Indira Hewlett,ⁿ Thomas N. Denny,^{c,d} Michael P. Busch,^{a,b} on behalf of the EQAPOL Program

^aBlood Systems Research Institute, San Francisco, California, USA

^bDepartment of Laboratory Medicine, University of California, San Francisco, California, USA

^cDuke Human Vaccine Institute, Duke University Medical Center, Durham, North Carolina, USA

^dDepartment of Medicine, Duke University Medical Center, Durham, North Carolina, USA

^eHenry M. Jackson Foundation, Silver Spring, Maryland, USA

^fMHRP, Walter Reed Army Institute of Research (WRAIR), Silver Spring, Maryland, USA

^gDepartment of Laboratory Medicine, University of Washington, Seattle, Washington, USA

^hBio-Rad Laboratories, Redmond, Washington, USA

ⁱOrtho Clinical Diagnostics, Rochester, New York, USA

^jHologic, Inc., San Diego, California, USA

^kQuanterix, Lexington, Massachusetts, USA

^lMesoScale Discovery, Rockville, Maryland, USA

^mPublic Health England, London, United Kingdom

ⁿCenter for Biologic Evaluation and Research (CBER), U.S. Food and Drug Administration (FDA), Bethesda, Maryland, USA

ABSTRACT Detection of acute HIV infection is critical for HIV public health and diagnostics. Clinical fourth-generation antigen (Ag)/antibody (Ab) combination (combo) and p24 Ag immunoassays have enhanced detection of acute infection compared to Ab-alone assays but require ongoing evaluation with currently circulating diverse subtypes. Genetically and geographically diverse HIV clinical isolates were used to assess clinical HIV diagnostic, blood screening, and next-generation assays. Three-hundred-member panels of 20 serially diluted well-characterized antibody-negative HIV isolates for which the researchers were blind to the results (blind panels) were distributed to manufacturers and end-user labs to assess the relative analytic sensitivity of currently approved and preapproved clinical HIV fourth-generation Ag/Ab combo or p24 Ag-alone immunoassays for the detection of diverse subtypes. The limits of detection (LODs) of virus were estimated for different subtypes relative to confirmed viral loads. Analysis of immunoassay sensitivity was benchmarked against confirmed viral load measurements on the blind panel. On the basis of the proportion of positive results on 300 observations, all Ag/Ab combo and standard sensitivity p24 Ag assays performed similarly and within half-log LODs, illustrating the similar breadth of reactivity and diagnostic utility. Ultrasensitive p24 Ag assays achieved dramatically increased sensitivities, while the rapid combo assays performed poorly. The similar performance of the different commercially available fourth-generation assays on diverse subtypes supports their use in broad geographic settings with locally circulating HIV clades and recombinant strains. Next-generation preclinical ultrasensitive p24 Ag assays achieved dramatically improved sensitivity, while rapid fourth-generation assays performed poorly for p24 Ag detection.

Received 2 January 2018 Returned for modification 28 January 2018 Accepted 4 April 2018

Accepted manuscript posted online 23 May 2018

Citation Stone M, Bainbridge J, Sanchez AM, Keating SM, Pappas A, Rountree W, Todd C, Bakkour S, Manak M, Peel SA, Coombs RW, Ramos EM, Shriver MK, Contestable P, Nair SV, Wilson DH, Stengelin M, Murphy G, Hewlett I, Denny TN, Busch MP, on behalf of the EQAPOL Program. 2018. Comparison of detection limits of fourth- and fifth-generation combination HIV antigen-antibody, p24 antigen, and viral load assays on diverse HIV isolates. *J Clin Microbiol* 56:e02045-17. <https://doi.org/10.1128/JCM.02045-17>.

Editor Angela M. Caliendo, Rhode Island Hospital

Copyright © 2018 American Society for Microbiology. All Rights Reserved.

Address correspondence to Mars Stone, mstone@bloodsystems.org.

KEYWORDS HIV, diagnostics, Ag/Ab combo, ultrasensitive Ag, viral load, analytic sensitivity, Ag detection

Detection of HIV in acute infection is critical for many aspects of HIV public health and clinical diagnostics (1). Diagnosis and early treatment to limit viral replication and establishment of latent reservoirs (2, 3) are vital. Additionally, the greatest risk of secondary transmission occurs during acute infection, when the viral load (VL) is high and HIV-specific binding and neutralizing antibodies (Abs) have not yet been formed (4). Thus, treatment and transmission intervention are optimal during the preseroconversion and early postseroconversion window periods, defined as the intervals between infection and reactivity by routine molecular and serological screening assays (5–8).

Fourth-generation p24 antigen (Ag)/Ab combination (combo) immunoassays detecting HIV p24 Ag, including several assays that discriminate p24 Ag and Ab reactivity, have narrowed the diagnostic window period from 3 to 4 weeks with Ab-only third-generation assays to within 2 weeks from the time of infection (2, 5, 9–11). This significantly enhances the detection of acute infection and reduces the transmission risk when the VL is the highest and the recently acquired and disseminating transmitted founder variants are most infectious (2, 4, 12). This is highlighted by recent updates to the Centers for Disease Control and Prevention (CDC) HIV testing guidelines, which recommend testing for established and acute infection with FDA-approved antigen/antibody combination immunoassays (13).

The rapid molecular evolution of HIV and the changing global subtype diversity challenge the accuracy and reliability of detection methods (12, 14), posing serious challenges for blood donor screening, surveillance, clinical diagnosis, and management domestically and internationally (14, 15), as these assays generally target specific genes or proteins, necessitating evaluation of different platforms across subtypes.

Thus, it is important to understand their comparative sensitivities and their ability to detect acute infection by performing ongoing evaluations of approved commercially available and preapproved assays (5, 16). The HIV seroconversion panels used in the development and performance evaluation of most current fourth-generation Ag/Ab combo assays were created over a decade ago, mainly consisted of subtype B samples, and are now outdated and thus not reflective of the currently circulating variants and the increasing proportions of non-B subtype or circulating and unique recombinant viruses (17). As these subtype-specific differences can affect assay performance, ongoing validation using clinically relevant contemporary viruses is essential (18).

To address this, well-characterized genetically diverse HIV clinical isolates assembled by the External Quality Assurance Program Oversight Laboratory (EQAPOL) were used to develop limit-of-detection (LOD) panels for assessment of the analytic sensitivities across platforms using p24 Ag-capture components, i.e., HIV fourth-generation diagnostic assays, stand-alone p24 Ag enzyme immunoassays (EIAs), novel p24 Ag digital detection technologies, and rapid point-of-care (POC) p24 Ag and Ab assays, relative to the analytic sensitivities of VL assays. Twenty isolates were selected from the EQAPOL Genotype Diversity Panel (19) to represent globally relevant subtypes and circulating recombinant forms (CRFs) and further characterized to confirm VL and p24 concentrations to ensure that the Ag/RNA ratios were consistent with the expected values (see Table S1 in the supplemental material). Both manufacturers and clinical and research labs participated in the comparison study (Table 1). The goals of the LOD study included (i) determination of the analytic sensitivities of currently approved and novel fourth-generation, laboratory, and rapid POC p24 Ag/Ab assays on divergent isolates, (ii) evaluation of the LOD and quantitation of novel ultrasensitive p24 detection assays, and (iii) demonstration of lab proficiency in detecting diverse isolates.

MATERIALS AND METHODS

Sample selection and panel construction. The two-stage panel distribution included (i) an initial LOD performance panel of 300 samples, including five-step 10-fold serial dilutions in triplicate of 20 well-characterized viral isolates, and (ii) a 20-sample panel to assess assay specificity that was distributed

TABLE 1 Participant sites and assays

| Site | Assay(s) run |
|--|---|
| Abbott | Architect HIV Ag/Ab combo EIA |
| Ortho Clinical Diagnostics | Vitros HIV combo assay |
| Bio-Rad | Bio-Rad GS HIV combo Ag/Ab EIA, Bio-Rad Genscreen HIV-1 p24 Ag assay |
| Walter Reed Army Institute of Research | Bio-Rad GS HIV combo Ag/Ab EIA, Bio-Rad Genscreen HIV-1 p24 Ag assay, Alere Determine HIV-1/2 Ag/Ab combo, SD Bioline HIV Ag/Ab combo assay (RUO) |
| University of Washington | Architect HIV Ag/Ab combo EIA, Abbott RealTime HIV viral load assay |
| Public Health England | Bio-Rad BioPlex 2200 HIV Ag-Ab combo assay |
| EQAPOL IVQAC ^a | PerkinElmer p24 Ag assay |
| MSD | Ultrasensitive p24 Ag assay |
| Quanterix | Ultrasensitive Simoa p24 Ag assay |
| FDA | Alere Determine HIV-1/2 Ag/Ab combo assay |
| Hologic | Aptima HIV-1 Quant Dx VL assay |

^aIVQAC, The Immunology Virology Quality Assessment Center.

to labs reporting positive results for the LOD panel. This panel included 10 negative controls and serial dilutions of selected clade B and clade C isolates from the LOD panel. Samples were selected as being representative from geographically diverse origins as a subset of the EQAPOL Genetic Diversity Panel and characterized by nearly full-length genome sequencing for subtype and recombination pattern by single-genome amplification (SGA) at the EQAPOL core laboratory and phylogenetic analysis, as previously described (see Fig. S1 in the supplemental material) (19). This genetic characterization confirmed that *gag* region sequences were consistent with the genotype assignment (see the GenBank accession number assignments in reference 19). Isolates with a ratio of VL/p24 within the expected range of 20×10^3 to 60×10^3 copies/pg were selected to ensure that the p24 Ag reactivity was attributable to virion-associated p24 Ag (Table S1) (19, 20). Additional spiking experiments were conducted to establish that lysed cell debris from infected culture cells would have made an insignificant contribution to supernatant VL and p24 levels (see Fig. S2 and the Fig. S2 legend).

High-titer viral culture supernatants were generated from preseroconversion (RNA-positive/Ab-negative) or recently acquired Ab-positive infected subject plasma samples that had been collected with ethical review and with approval of the collection protocol and informed consent as previously established for the donor organization's collection activities (20). Supernatants were serially diluted to concentrations of $\sim 10^6$ to 10^2 copies/ml (cp/ml) in defibrinated human serum (Gemini Biosciences), frozen into 1-ml aliquots at -80°C , and assembled into panels for which the researchers were blind to the results (referred to here as blind panels). Prior to making the final serial dilutions of samples to be included in the panel, the supernatant concentrations of the 20 selected isolates were quantified by the Cobas AmpliPrep/Cobas TaqMan HIV-1 test (v2.0) at an intermediate dilution 1 log above the highest concentration to be included in the panel. To confirm the viral load concentrations of the final dilutions, the full blind panel was retested (by the Abbott RealTime HIV-1 viral load assay and Hologic Aptima HIV-1 Quant Dx VL assay).

Assays performed. To assess the performance of the viral load Ag/Ab combo and Ag-alone assays, identical panels were tested by 10 platforms at 11 sites. Viral load assays included the Hologic Aptima HIV-1 Quant Dx VL (21) and Abbott RealTime HIV-1 VL assays (22). HIV Ag/Ab combo fourth-generation assays included the Abbott Architect HIV Ag/Ab combo EIA (23), Ortho Clinical Diagnostics Vitros HIV combo assay, Bio-Rad GS HIV combo Ag/Ab EIA (catalog no. 26218), and Bio-Rad BioPlex 2200 HIV Ag-Ab combo assay. We also evaluated the Bio-Rad Genscreen HIV-1 p24 Ag assay (catalog no. 71120). Ultrasensitive preclinical p24 assays included the MesoScale Discovery (MSD) p24 Ag (24) and Quanterix Simoa p24 Ag (25) assays. Rapid POC immunochromatographic assays included the Alere Determine HIV-1/2 Ag/Ab combo and SD Bioline HIV Ag/Ab combo assays. Three assay platforms (the Bio-Rad GS HIV combo Ag/Ab EIA, Abbott Architect HIV Ag/Ab combo EIA, and Bio-Rad Genscreen HIV-1 p24 Ag platforms) were tested by both the manufacturer and an end user clinical lab. The participating sites and assays performed are summarized in Table 1. The assays were performed according to the manufacturers' package insert specifications and summarized assay methods, and the basis for cutoff determinations is included in the supplemental material.

Viral load comparison and Bland-Altman plot. Methods for graphically comparing VL assays (the Hologic and Abbott assays) included a scatterplot comparing the results of the two assays and a Bland-Altman plot where the Hologic Aptima assay was designated the reference assay (26).

Logistic modeling process. A logistic model was created with detection above the cutoff set by the manufacturer using the outcome (0 = no detection, 1 = detection) for each sample. The independent variables considered were sample VL (\log_{10} scale), assay type, virus subtype, and the interaction of assay and virus subtype. A forwards and backwards modeling process was used with a criterion of a Wald's chi-square *P* value of <0.01 for parameter inclusion. A main effects logistic model with parameters for the \log_{10} VL, assay, and virus subtype was fit. In the second model, the interaction of assay and virus subtype was included. Given the significance of all the terms ($P < 0.001$) in the main effects model and the lack of significance of the interaction term, the modeling process was completed. Graphics showing the probability of virus detection for the different assays and subtypes relative to the \log_{10} viral load based upon the main effects model were generated. In the first analysis, the functions of the different assays using subtype C as the reference are shown. In the second analysis, the functions of the different subtypes using the Bio-Rad GS HIV combo Ag/Ab EIA as the reference are shown.

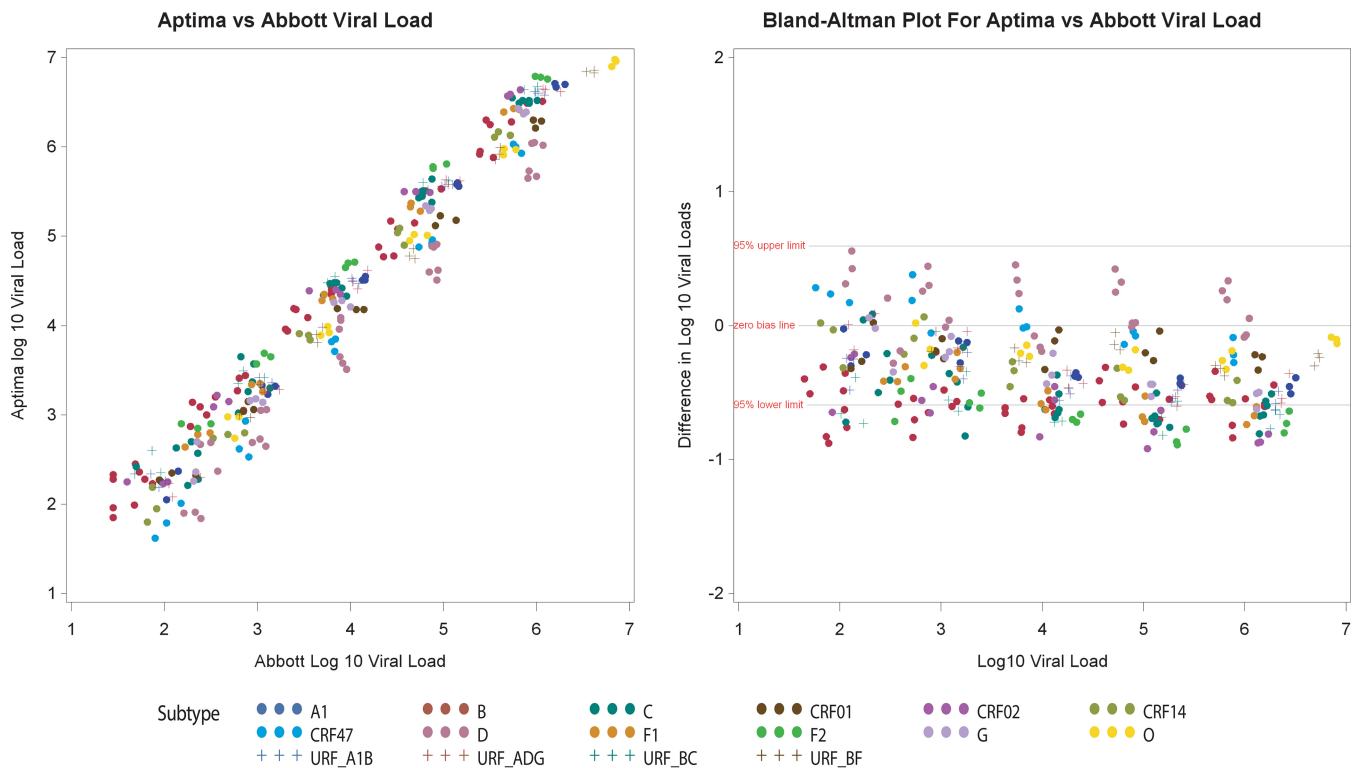


FIG 1 Comparison and confirmation of viral load. (Left) Correlation between Abbott RealTime HIV viral load assay results (performed by the University of Washington) and the Hologic Aptima HIV-1 Quant Dx VL assay results (Aptima viral load). (Right) The difference in viral loads between the Aptima and Abbott assays. Each dot represents a sample, with the total number being 300. The color denotes the HIV subtype of the sample. Pearson's r correlation was determined to be 0.98.

Demonstrated viral load at the manufacturer's cutoff. The logistic regression model used to estimate the VL for each assay at the manufacturer's positivity threshold value used maximum likelihood and included a separate cubic trajectory of the signal-to-cutoff (S/CO) ratio value for each lab/assay combination. Since the S/CO values were log transformed and, thus, the threshold value was zero, this permitted simply using the intercept parameter and the confidence intervals for each laboratory/assay combination as the estimator for the VL at the assay threshold.

RESULTS

Confirmation and comparison of viral load assays with the LOD panel. To confirm the concentrations of serially diluted samples and compare the performance of quantitative FDA-approved VL assays, the full blind panel was tested by the Abbott RealTime HIV VL assay and the Hologic Aptima HIV-1 Quant Dx VL assay. The Aptima and Abbott VL data correlated well, with a Pearson's r of 0.98 (Fig. 1). There appeared to be a mean shift which varied by subtype, with the Aptima assay giving 0.36 \log_{10} higher values, on average (Fig. 1, left), with an apparent downward trend toward greater differences at higher concentrations (Fig. 1, right); however, the results were within the expected assay variability.

Proportion of positive results of p24 and HIV Ag/Ab combo assays with the LOD panel. The proportion of positive results for each lab was based on the percentage of 300 samples in the initial performance panel reported positive by each assay (Table 2). Qualitatively, the assay results clustered into three tiers: (i) the next-generation preclinical ultrasensitive p24 assays from MSD and Quanterix performed with the highest sensitivity, with ~70 to 99% of the samples being reactive; (ii) the FDA-approved Ag/Ab combo and CE-marked (approved for use by the Council of Europe) stand-alone p24 assays all performed similarly, with 31 to 40% of samples being reactive; and (iii) rapid detection assays yielded negative results with all samples. Table 2 also includes the estimated average VL across subtypes at the manufacturers' cutoffs for assays reporting reactive results for p24 Ag detection. These VLs ranged from

TABLE 2 Proportion of positive results and VLs at the cutoff by assay and lab^a

| Assay category | Assay | Lab | Proportion positive on performance panel (n = 300) | Estimated VL at the manufacturer's cutoff ^b | Proportion positive on specificity panel (n = 20) |
|--|---|-----------------|--|--|---|
| Fourth-generation HIV Ag/Ab | Bio-Rad GS HIV combo Ag/Ab EIA | BR | 0.40 | 4.71 (4.61, 4.80) | 0 |
| | Abbott Architect HIV Ag/Ab combo EIA | WR | 0.33 | 4.88 (4.79, 4.97) | ND |
| | | ABT | 0.37 | 4.70 (4.62, 4.79) | 0 |
| | | UW | 0.37 | 4.60 (4.52, 4.69) | ND |
| | Vitros HIV combo assay | OCD | 0.32 | 5.06 (4.96, 5.15) | 0 |
| Bio-Rad BioPlex 2200 HIV Ag-Ab combo assay | PHE | 0.31 | 4.91 (4.77, 5.06) | ND | |
| HIV p24 | Bio-Rad Genscreen HIV-1 Ag assay | BR | 0.40 | 4.41 (4.32, 4.50) | 0 |
| | | WR | 0.40 | 4.60 (4.50, 4.69) | ND |
| Ultrasensitive HIV p24 | Quanterix ultrasensitive Simoa p24 assay | QT | 0.99 | 2.34 (1.45, 4.15) ^c | 0 |
| | | MSD | 0.69 | 2.68 (1.45, 3.77) ^c | 0 |
| POC rapid assays | Alere Determine HIV-1/2 Ag/Ab combo assay | FDA | 0 | NA | ND |
| | | WR ^d | 0 | NA | ND |
| | | WR ^d | 0 | NA | ND |
| | SD Bionline HIV Ag/Ab combo assay | WR ^d | 0 | NA | ND |

^aABT, Abbott; UW, University of Washington; BR, Bio-Rad; OCD, Ortho Clinical Diagnostics; PHE, Public Health England; MSD, MesoScale Discovery; QT, Quanterix; FDA, U.S. Food and Drug Administration; WR, Walter Reed Army Institute of Research; ND, not determined; NA, not applicable.

^bThe values represent the viral load (95% CI), unless indicated otherwise. The VL (log₁₀ number of copies per milliliter) is based on quantification of the blind panel tested by the Hologic Aptima HIV Quant Dx assay.

^cPreclinical assays; not FDA approved. The values represent the median (minimum, maximum).

^dThe first 243 samples were negative or invalid; testing was discontinued.

4.41 log₁₀ cp/ml (95% confidence interval [CI], 4.32, 4.50 log₁₀ cp/ml) for the Bio-Rad Genscreen HIV-1 p24 Ag assay to 5.06 log₁₀ cp/ml (95% CI, 4.96, 5.15 log₁₀ cp/ml) for the Vitros HIV combo assay. For the ultrasensitive p24 Ag assays, the sensitivity was reported as the median and minimum-maximum VLs at the manufacturer's cutoff for all subtypes, due to outliers (see Fig. 6a and b), which violate the model assumptions and would lead to spurious results if they were included in the model.

Predicted proportion of viral subtype C isolates detected across Ag/Ab combo and Ag assays. Figure 2 (left) presents the model-based predicted VL at the cutoff for samples detectable by each assay utilizing HIV subtype C to illustrate the relative rank order performance with reactive S/CO by VL among the four FDA-approved Ag/Ab combo assays and one CE-marked p24 Ag assay. The Bio-Rad GS combo Ag/Ab EIA, Bio-Rad Genscreen HIV-1 p24 Ag assay and Abbott Architect Ag/Ab combo EIA coupled closely, detecting 50% of panel samples with VLs of between 4.4 and 4.6 log₁₀ copies/ml. The Ortho Vitros HIV combo and BioPlex 2200 HIV Ag-Ab assays had similar

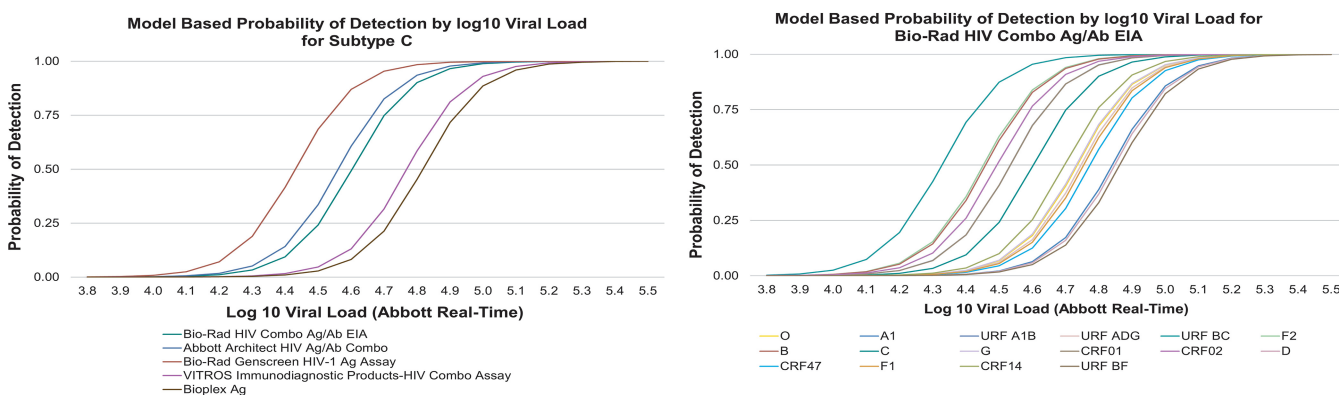


FIG 2 Model-based proportion detected by assay for subtype C and by viral subtype detected by the Bio-Rad GS HIV combo Ag/Ab EIA. (Left) Predicted proportion of samples detectable by each of the four FDA-approved Ag/Ab combo assays and the one CE-marked p24 Ag-alone assay utilizing HIV subtype C to illustrate the relative rank order of performance with reactive S/CO by VL. (Right) Predicted proportion of viral subtypes detectable by the Bio-Rad GS HIV combo Ag/Ab EIA, which was used as a representative of the relative rank order performance with reactive S/CO by VL.

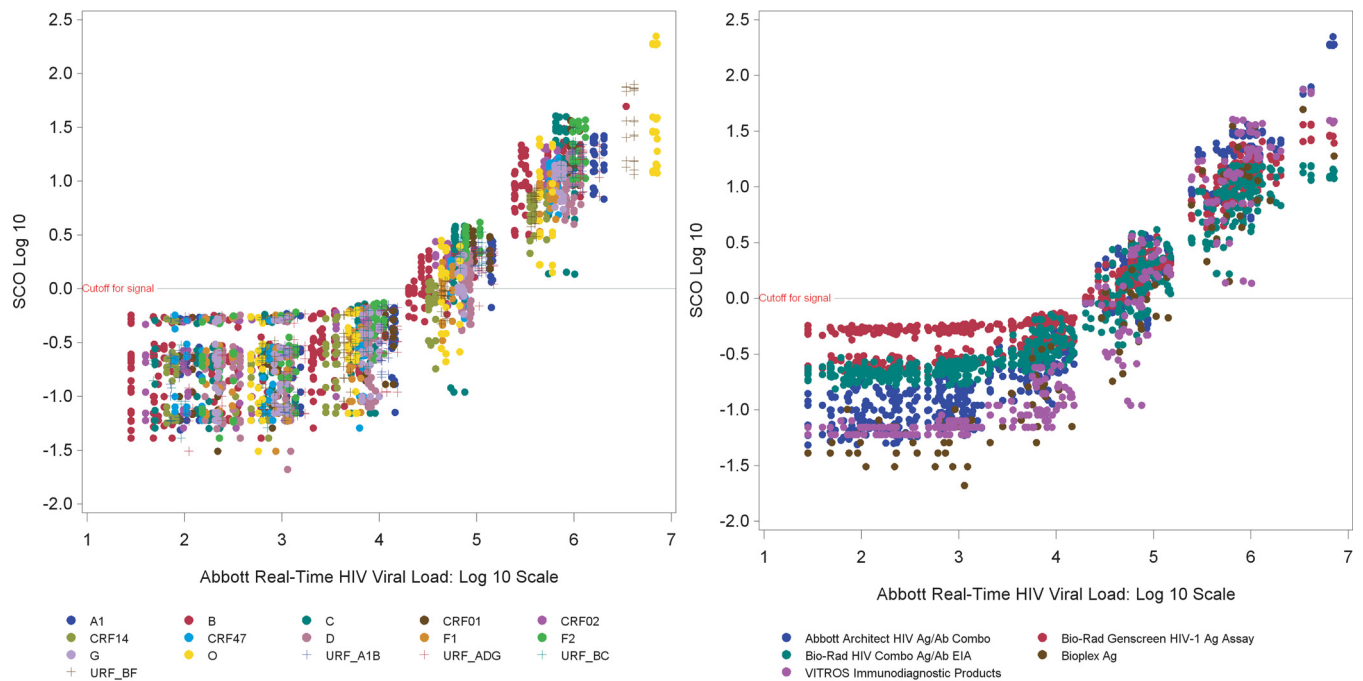


FIG 3 All S/CO data by viral load. S/CO data by viral load were based on viral subtypes (left) and assay platform (right).

performance, detecting 50% of samples with VLs of between 4.7 and 4.8 \log_{10} copies/ml. Although there were small differences, only the extremes of proportions detected by the assays likely have biologically relevant differences, as the 50% LODs of all assays are clustered between 4.4 and 4.8 \log_{10} copies/ml and a half-log difference in VLs was required to infer clinically meaningful changes in VLs. Subtype C was chosen as a reference category, as it is a prevalent circulating subtype and was represented by two samples in the panel, providing informative subtype replicates. The assays had a similar relative rank order performance for detection of the other subtypes.

Predicted proportion of viral subtypes detected. The Bio-Rad GS HIV Ag/Ab EIA results are shown in Fig. 2 (right) as an illustrative example to present representative sensitivities across subtypes with our model analysis. Among the subtypes tested, the p24 Ag of the URF-BF, A1, and D subtypes was the most difficult to detect, and p24 of recombinant subtype URF-BC was detected at the highest frequency. Other fourth-generation assays yielded similar findings and had the same rank order of detection across diverse isolates. The breadth of reactivity, i.e., the difference in detection among subtypes, was similar among all assays, with no evidence that one platform had significantly better detection among the different subtypes by these models.

Quantitative results. S/CO data were used to refine the proportion of samples detected to infer the LOD for the HIV p24 Ag signal relative to the VL (Table 2 and Fig. 3). Overall, HIV Ag/Ab combo and Ag-alone assays demonstrated similar performance with diverse subtypes. There was evidence for small differences in detection by subtype by the Abbott Architect HIV Ag/Ab combo EIA for subtype CRF02, the Vitros HIV combo assay for subtype C, and the MSD assay for subtypes CRF01 and URF-BF, but these differences were likely not clinically meaningful. The Quanterix assay appeared to have the greatest outlier issue for the detection of subtype O virus (see Fig. 6, right), although to maintain a logistically manageable panel size, each of these subtypes was represented by a single isolate in the panel, and it is possible that these differences were due to sample variability rather than differential subtype detection.

Due to logistical constraints, only 100 samples representing each isolate in singlet were run on the Bio-Rad BioPlex platform, but the assay appeared to perform similarly for all subtypes (Fig. 4a).

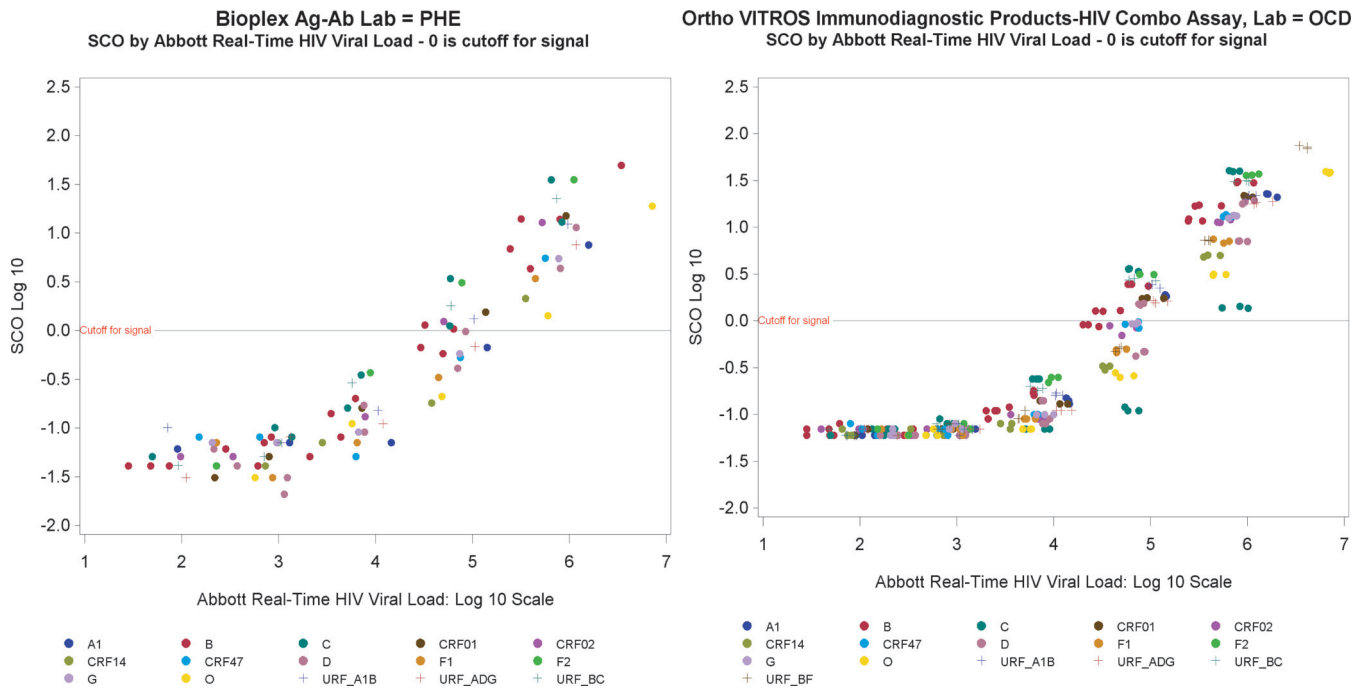


FIG 4 S/CO relative to viral load for the BioPlex 2200 HIV Ag-Ab assays and the Ortho Clinical Diagnostics Vitros HIV combo assay. (Left) Results of the Bio-Rad 2200 HIV BioPlex Ag-Ab assay run at Public Health England (PHE); (right) results of the Vitros HIV combo assay run at Ortho Clinical Diagnostics (OCD).

Comparison of performance by manufacturer and clinical lab. Blind panels were tested in parallel by the manufacturer and an end user lab on three platforms: the Abbott Architect HIV Ag/Ab combo, the Bio-Rad GS HIV combo Ag/Ab EIA, and Bio-Rad Genscreen HIV-1 p24 Ag assay platforms. The Abbott Architect assay performed similarly when it was performed by the manufacturer and the clinical lab on the basis of both the percentage of reactive samples (0.37) and quantitative analysis of S/CO and VL values to yield LODs (Fig. 5, left). The demonstrated VL at the cutoff across all subtypes was similar for the manufacturer and the clinical lab, with values of 4.70 log₁₀ cp/ml (95% CI, 4.62, 4.79 log₁₀ cp/ml) and 4.60 log₁₀ cp/ml (95% CI, 4.52, 4.69 log₁₀ cp/ml), respectively. The Bio-Rad GS HIV combo Ag/Ab EIA had a slightly higher proportion of positive results when it was performed by the manufacturer (0.40) than when it was performed by the clinical lab (0.33), and there was a small but statistically significant

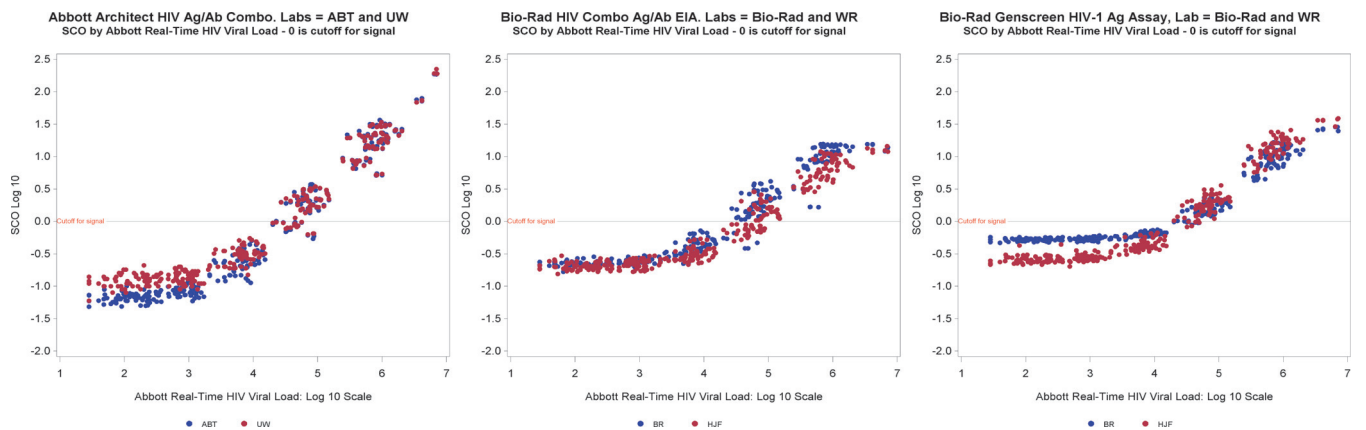


FIG 5 Comparing assays run in both manufacturer and clinical labs. (Left) The Abbott Architect HIV Ag/Ab combo assay was performed at the University of Washington (UW) and at Abbott (ABT; the manufacturer's lab). (Middle) The Bio-Rad GS HIV combo Ag/Ab EIA was performed at Bio-Rad (the manufacturer's lab) and at WRAIR (WR). (Right) The Bio-Rad Genscreen HIV-1 p24 Ag assay was performed at WRAIR and at Bio-Rad (the manufacturer's lab). Blue, the manufacturer's lab; red, the clinical lab.

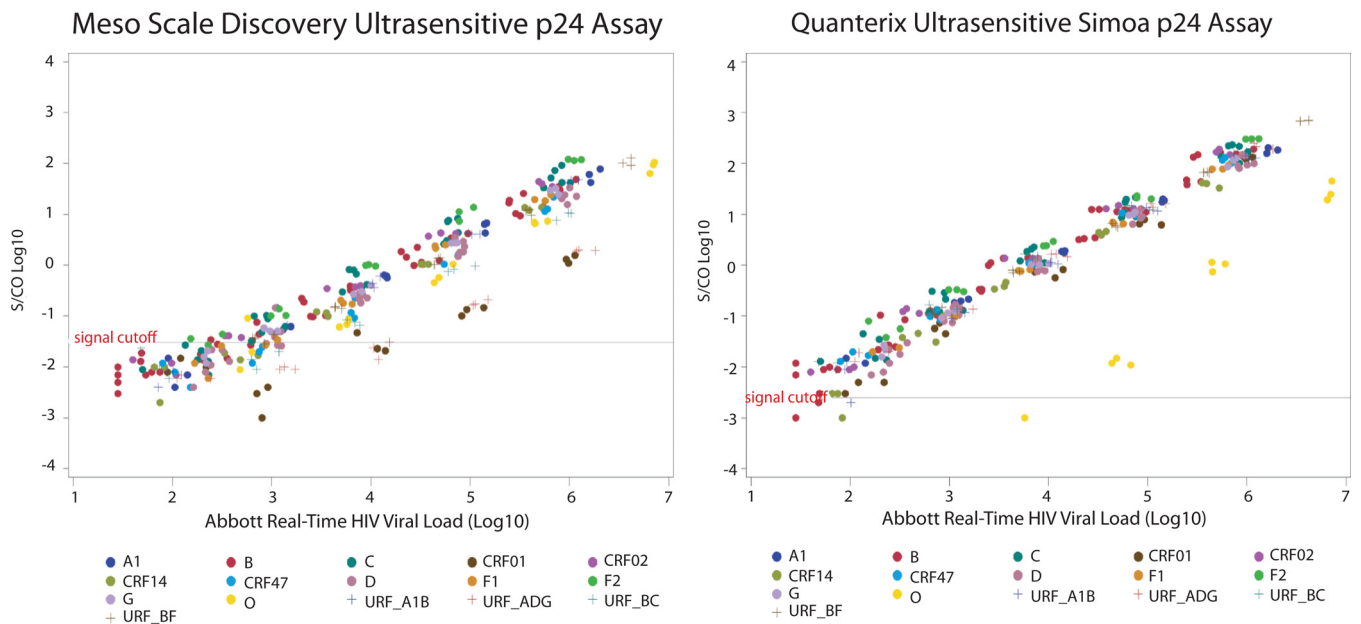


FIG 6 Ultrasensitive p24 Ag assay signal intensity compared to viral load. (Left) MSD ultrasensitive p24 assay; (right) Quanterix Ultrasensitive Simoa assay.

difference in the demonstrated VL at the cutoff, with the manufacturer providing a VL at the cutoff of 4.71 \log_{10} cp/ml (95% CI, 4.61, 4.80 \log_{10} cp/ml) and the clinical lab providing a VL at the cutoff of 4.88 \log_{10} cp/ml (95% CI, 4.79, 4.97 \log_{10} cp/ml). The Bio-Rad Genscreen HIV-1 p24 Ag assay performed slightly better when it was run by the manufacturer than when it was run by the clinical lab, on the basis of the percentage of reactive samples (0.40 versus 0.33), although quantitative analysis of the S/CO relative to the VL showed a tighter distribution across all samples with a plateau in the signal at a higher S/CO (Fig. 5, right). This was observed in the signal below the manufacturer's cutoff and likely does not affect the interpretation of positivity.

Next-generation HIV Ag assays. The Quanterix Simoa p24 and MSD p24 Ag assay results (Table 2 and Fig. 6) illustrated the relationship between the VL and the Ag signal intensity and supported the conclusion that these newer-generation preclinical p24 Ag detection platforms achieved enhanced sensitivity relative to the currently approved fourth-generation Ag/Ab combo and stand-alone p24 Ag assays.

POC assays. The rapid detection assays Alere Determine HIV-1/2 Ag/Ab combo and SD Bioline HIV Ag/Ab combo yielded no positive results, and testing on the Bioline platform was discontinued after 243 negative results. Follow-up studies were conducted at two participating laboratories, which demonstrated that this was not due to inhibition of the diluent since virus was readily detected by both assays in the samples with higher virus concentrations used to produce the panel.

Performance of specificity panel. Assays returning positive results with the performance LOD panel were reassessed for specificity using a second panel. All assays evaluated demonstrated excellent specificity, with no false-positive results being reported (Table 2).

DISCUSSION

The EQAPOL Viral Diversity Program was established to develop repositories for validation of diagnostic tools relevant to the evolving HIV genetic diversity, including newly circulating recombinant forms (CRFs) and drug resistance mutations (20). These well-characterized isolates of dynamic global HIV populations enabled rigorous evaluation and comparative analysis of the LODs of the p24 Ag detection components of fourth-generation Ag/Ab combo assays relative to those of stand-alone HIV p24 Ag assays. These are relevant for diagnostic and blood donor screening settings, where

testing relies on the sensitivity of assays to identify new infections that are within the Ab-negative window period of acute viremia.

Confirmation of the concentrations of serially diluted HIV RNA on multiple VL platforms allowed a comparison of the performance of quantitative viral load assays from Abbott and Hologic, both of which are FDA-approved assays with excellent sensitivities and well-documented performance with genotype variants (27–30). The results demonstrated excellent correlations across the range of viral loads for all subtypes for both assays, validated the concentrations of HIV RNA in the panel itself, and provided additional evidence of the comparability of the results when the assays were performed on different p24 Ag detection platforms.

The commercially available p24 Ag/Ab and Ag stand-alone assays were demonstrated to have similar sensitivities across most clades and recombinant forms. The rank order of detection of different subtypes using comparable assays was consistent, with similar relative levels of detection across subtypes. There was evidence for differences in subtype detection for some assays, but the differences were small. The sequences within *gag* were consistent with the genotype assignment for each isolate, and no mutations that could explain differential subtype detection were detected (see the GenBank accession numbers in reference 19 for access to *gag* sequences); furthermore, small sequence differences likely do not affect the p24 tertiary structure and, hence, the capacity of monoclonal Abs to detect p24 Ag binding in immunoassays. Due to logistical constraints in the number of samples distributed in the panels, most of the subtypes were represented by a single isolate in the panel; consequently, it is possible that these differences were due to sample variation rather than systematic differences in subtype detection, although if this were the case, these sample differences would likely have been demonstrated across all assays.

For the 300 samples tested, 31 to 40% were detectable with overlapping confidence intervals, indicating no significant difference in performance among the assays. When comparing the performance of Ag/Ab combo assays to that of Ag assays, the proportion of positive results was nearly identical according to the range of positive results, indicating that the addition of Ag detection to Ab detection did not impact assay performance. When analyzed to estimate LODs, all FDA-approved Ag/Ab combo and Ag assays had detectable signal within a half-log VL of each other, illustrating the similar diagnostic utilities of all assays. Considering the diversity of the panel, the similar performance across platforms demonstrates the utility of these assays in broad geographic settings.

Although samples were evaluated for positivity on the basis of the manufacturers' reported S/CO cutoff, we found that several Ag/Ab combo and Ag assays had a dynamic range of signal and LODs that were lower than the manufacturers' cutoff values. For example, the Vitros, Architect, and Bio-Rad combo assays appeared to have the potential for enhanced sensitivity with a continued dynamic range below the labeled S/CO threshold. The Vitros and Abbott Architect assays demonstrated a signal dynamic range and a calculated limit of blanks ~ 1 log below the manufacturers' S/CO cutoff. To a lesser degree, the Bio-Rad Genscreen HIV-1 p24 Ag assay also had a capability of detection lower than that currently being utilized. This suggests that the signal detected below the FDA's assigned cutoff for licensed assays could enable detection of viremia earlier in infection. The manufacturer's claimed LODs, based on FDA criteria, may be more conservative than the assay is capable of detecting, given the necessity to balance LOD sensitivity with avoiding false-positive results and preserving optimal specificity.

Performance reproducibility among end users was assessed through parallel testing by the manufacturer and a research lab by the Abbott Architect HIV Ag/Ab combo, Bio-Rad GS HIV Ag/Ab combo, and Bio-Rad Genscreen HIV-1 p24 Ag assays (Fig. 5). The proportions of positive results returned by the manufacturer and testing lab were identical for the Abbott Architect HIV Ag/Ab combo assay. The Bio-Rad GS HIV Ag/Ab combo assay performed by the manufacturer showed a slightly higher proportion of positive results and a lower VL at the cutoff (Table 2) but a similar distribution of signals

across all samples (Fig. 5b). The disparity in the detection by the Bio-Rad GS HIV Ag/Ab combo assay between the manufacturer and lab, although it minimally reached the level of significance, suggests that differences or inconsistencies in assay execution may affect assay performance.

Next-generation preclinical ultrasensitive HIV Ag detection assays, such as the MSD and Quanterix Simoa digital p24 Ag assays, achieved sensitivities dramatically better than those of the current commercially available assays. These assays could provide a mechanism to detect acute HIV infection with minimal sample handling and a short turnaround time to best facilitate patient intervention. Of note, the Quanterix Simoa p24 Ag immunoassay does not perform immune complex dissociation; while this is not pertinent to the performance of the assays using this panel of serially diluted RNA-positive/Ab-negative plasma-derived culture supernatants, testing of clinical samples for VL would require the addition of an Ab dissociation step, potentially impacting the performance demonstrated here. However, simply testing clinical samples for early HIV infection would not require an Ab dissociation step, as is the case with the combo assays, wherein early detection is accomplished with p24 reactivity, while chronic infection is detected through reactivity for HIV antibodies.

The immunochromatographic rapid tests from Bioline and Alere did not perform well with this panel. Our data show that the Alere Determine HIV-1/2 Ag/Ab combo and SD Bioline HIV Ag/Ab combo assays were less sensitive for the detection of HIV p24 than the other laboratory-based Ag/Ab combo and Ag assays tested. We cannot say that these POC assays are not capable of detecting p24 Ag in plasma during acute infections, since we only tested dilutions of cultured virus and not early infection clinical specimens. In fact, the Alere product insert shows that the Alere (p24 Ag component) assay detected 19 of the 28 seroconversion panels earlier than an FDA-licensed anti-HIV-1/2 EIA (31). Another study of serum samples reported sensitivities of the Alere assay of 95% and the SD Bioline assay of 91%; the p24 Ag component of the Alere test detected 10 of 13 acute infection samples (RNA positive, antibody negative), while the SD Bioline assay detected only 1 (32). However, in our study, these assays did not detect samples from the panel at any dilution, indicating either that the Ag component of these assays is not sensitive enough to detect the concentrations tested or that there was inhibition in the assay particular to the preparation of these panels. Further studies conducted at the Walter Reed Army Institute of Research (WRAIR) and FDA to determine if there was inhibition demonstrated detection of representative diverse HIV isolates in our panel at higher concentrations and no inhibition by the diluent used to construct the original study panels. Of note, the Alere Determine HIV-1/2 Ag/Ab combo assay that was evaluated in this study is the FDA-approved test distributed in the United States, but there is an enhanced Alere HIV combo rapid test that is CE marked and commercially available outside the United States that has demonstrated improved sensitivity for the detection of p24 Ag during acute HIV infection (33). POC assays provide a result at the time of testing and thus are broadly relied on for clinical diagnosis in field settings and, in particular, in resource-limited settings to identify HIV infection for treatment initiation and public health intervention. Although POC tests enable expanded access to testing, their limited performance in detecting acute HIV infection in clinical samples has been reported and must be considered (34, 35).

In conclusion, the development and comparative testing of diverse blind panels are essential to assess the relative performance of commercially available and preclinical assays. We confirmed similar performance across subtypes between currently FDA-approved fourth-generation and stand-alone p24 assays, demonstrated the enhanced sensitivity of next-generation p24 digital detection platforms, and observed the poor performance of POC assays. The EQAPOL diversity panel brings a unique sample set to the evaluation of diagnostic and research assays to allow unbiased assessment of performance and will be essential in the development and validation of new diagnostic tools for identifying future recombinant strains of HIV.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/JCM.02045-17>.

SUPPLEMENTAL FILE 1, PDF file, 0.3 MB.

ACKNOWLEDGMENTS

This study was supported by the National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH), for the External Quality Assurance Program Oversight Laboratory (EQAPOL) (HHSN272201700061C and HHSN272200800014C), the Global HIV panel project (DAIDS IAA-Y1-HB-5026-01) and DAIDS IAA 117023, the ACTG Laboratory Center (UM1-AI-106701), and the University of Washington Center for AIDS Research (P30-AI-027757).

We thank Mohan Haleyur Giri Setty for his technical assistance.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

M. Kathleen Shriver is an employee of Bio-Rad Laboratories, Paul Contestable is an employee of Ortho Clinical Diagnostics, Sangeetha Vijaysri Nair is an employee of Hologic, Inc., David H. Wilson is an employee of Quanterix, and Martin Stengelin is an employee of MesoScale Discovery.

This article reflects the views of the authors and should not be construed to represent the views or policies of the FDA.

REFERENCES

- Delwart EL, Kalmin ND, Jones TS, Ladd DJ, Foley B, Tobler LH, Tsui RCP, Busch MP. 2004. First report of human immunodeficiency virus transmission via an RNA-screened blood donation. *Vox Sang* 86:171–177. <https://doi.org/10.1111/j.0042-9007.2004.00416.x>.
- Hollingsworth TD, Anderson RM, Fraser C. 2008. HIV-1 transmission, by stage of infection. *J Infect Dis* 198:687–693. <https://doi.org/10.1086/590501>.
- Ananworanich J, Chomont N, Eller LA, Kroon E, Tovanabutra S, Bose M, Nau M, Fletcher JL, Tipsuk S, Vandergeeten C, O'Connell RJ, Pinyakorn S, Michael N, Phanuphak N, Robb ML, RV217 and RV254/SEARCH010 Study Groups. 2016. HIV DNA set point is rapidly established in acute HIV infection and dramatically reduced by early ART. *EBioMedicine* 11: 68–72. <https://doi.org/10.1016/j.ebiom.2016.07.024>.
- Ma Z-M, Stone M, Piatak M, Jr, Schweighardt B, Haigwood NL, Montefiori D, Lifson JD, Busch MP, Miller CJ. 2009. High specific infectivity of plasma virus from the pre-ramp-up and ramp-up stages of acute simian immunodeficiency virus infection. *J Virol* 83:3288–3297. <https://doi.org/10.1128/JVI.02423-08>.
- Busch MP, Glynn SA, Stramer SL, Strong DM, Caglioti S, Wright DJ, Pappalardo BL, Kleinman SH. 2005. A new strategy for estimating risks of transfusion-transmitted viral infections based on rates of detection of recently infected donors. *Transfusion* 45:254–264. <https://doi.org/10.1111/j.1537-2995.2004.04215.x>.
- Delaney KP, Hanson DL, Masciotra S, Ethridge SF, Wesolowski L, Owen SM. 2017. Time until emergence of HIV test reactivity following infection with HIV-1: implications for interpreting test results and retesting after exposure. *Clin Infect Dis* 64:53–59. <https://doi.org/10.1093/cid/ciw666>.
- Kleinman S, Busch MP, Korelitz JJ, Schreiber GB. 1997. The incidence/window period model and its use to assess the risk of transfusion-transmitted human immunodeficiency virus and hepatitis C virus infection. *Transfus Med Rev* 11:155–172. <https://doi.org/10.1053/tmrv.1997.0110155>.
- Masciotra S, McDougal JS, Feldman J, Sprinkle P, Wesolowski L, Owen SM. 2011. Evaluation of an alternative HIV diagnostic algorithm using specimens from seroconversion panels and persons with established HIV infections. *J Clin Virol* 52(Suppl 1):S17–S22. <https://doi.org/10.1016/j.jcv.2011.09.011>.
- Jurriaans S, Back NKT, Wolthers KC. 2011. Ten years of HIV testing with fourth generation assays: the Amsterdam experience. *J Clin Virol* 52(Suppl 1):S67–S69. <https://doi.org/10.1016/j.jcv.2011.09.021>.
- Powers KA, Ghani AC, Miller WC, Hoffman IF, Pettifor AE, Kamanga G, Martinson FEA, Cohen MS. 2011. The role of acute and early HIV infection in the spread of HIV and implications for transmission prevention strategies in Lilongwe, Malawi: a modelling study. *Lancet* 378:256–268. [https://doi.org/10.1016/S0140-6736\(11\)60842-8](https://doi.org/10.1016/S0140-6736(11)60842-8).
- Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Li X, Laeyendecker O, Kiwanuka N, Kigozi G, Kiddugavu M, Lutalo T, Nalugoda F, Wabwire-Mangen F, Meehan MP, Quinn TC. 2005. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *J Infect Dis* 191:1403–1409. <https://doi.org/10.1086/429411>.
- Weber B, Mbarigane Fall EH, Berger A, Doerr HW. 1998. Reduction of diagnostic window by new fourth-generation human immunodeficiency virus screening assays. *J Clin Microbiol* 36:2235–2239.
- Centers for Disease Control and Prevention. 2017. National HIV testing day and new testing recommendations. *MMWR Morb Mortal Wkly Rep* 63:537.
- Delwart E, Slikas E, Stramer SL, Kamel H, Kessler D, Krysztof D, Tobler LH, Carrick DM, Steele W, Todd D, Wright DJ, Kleinman SH, Busch MP, NHLBI-REDS-II Study Group. 2012. Genetic diversity of recently acquired and prevalent HIV, hepatitis B virus, and hepatitis C virus infections in US blood donors. *J Infect Dis* 205:875–885. <https://doi.org/10.1093/infdis/jir862>.
- Luft LM, Gill MJ, Church DL. 2011. HIV-1 viral diversity and its implications for viral load testing: review of current platforms. *Int J Infect Dis* 15:e661–e670. <https://doi.org/10.1016/j.ijid.2011.05.013>.
- Qiu X, Sokoll L, Yip P, Elliott DJ, Dua R, Mohr P, Wang XY, Spencer M, Swanson P, Dawson GJ, Hackett J. 2017. Comparative evaluation of three FDA-approved HIV Ag/Ab combination tests using a genetically diverse HIV panel and diagnostic specimens. *J Clin Virol* 92:62–68. <https://doi.org/10.1016/j.jcv.2017.05.005>.
- Hemelaar J, Gouws E, Ghys PD, Osmanov S, WHO-UNAIDS Network for HIV Isolation and Characterisation. 2011. Global trends in molecular epidemiology of HIV-1 during 2000–2007. *AIDS* 25:679–689. <https://doi.org/10.1097/QAD.0b013e328342ff93>.
- Antunes R, Figueiredo S, Bartolo I, Pinheiro M, Rosado L, Soares I, Lourenco H, Taveira N. 2003. Evaluation of the clinical sensitivities of three viral load assays with plasma samples from a pediatric population predominantly infected with human immunodeficiency virus type 1 subtype G and BG recombinant forms. *J Clin Microbiol* 41:3361–3367. <https://doi.org/10.1128/JCM.41.7.3361-3367.2003>.
- Hora B, Keating SM, Chen Y, Sanchez AM, Sabino E, Hunt G, Ledwaba J, Hackett J, Jr, Swanson P, Hewlett I, Ragupathy V, Vikram Vemula S, Zeng P, Tee K-K, Chow WZ, Ji H, Sandstrom P, Denny TN, Busch MP, Gao F, on

- behalf of the REDS-III and EQAPOL Programs. 2016. Genetic characterization of a panel of diverse HIV-1 isolates at seven international sites. *PLoS One* 11:e0157340. <https://doi.org/10.1371/journal.pone.0157340>.
20. Sanchez AM, Demarco CT, Hora B, Keinonen S, Chen Y, Brinkley C, Stone M, Tobler L, Keating S, Schito M, Busch MP, Gao F, Denny TN. 2014. Development of a contemporary globally diverse HIV viral panel by the EQAPOL Program. *J Immunol Methods* 409:117–130. <https://doi.org/10.1016/j.jim.2014.01.004>.
 21. Aptima Inc. 2015. HIV-1 Quant Dx assay (package insert). AW-11853 Rev 002. Aptima Inc, San Diego CA.
 22. Abbott Molecular Inc. Abbot RealTime HIV-1 assay (package insert). Abbott Molecular Inc, Des Plaines IA. <https://www.molecular.abbott/sal/en-us/staticAssets/realtime-hiv-1-package-insert.pdf>.
 23. Abbott Laboratories. 2009. Architect HIV Ag/Ab combo. Abbott Laboratories, Wiesbaden, Germany. <http://www.ilcxmedical.com/files/PDF/HIVAgAbcombo.pdf>.
 24. Stengelin M, Roy D, Aghvanyan A, Kenten J, Sigal GB, Glezer EN, Wohlstradter JN. 2015. HIV p24 immunoassay with the sensitivity of PCR methods. MesoScale Diagnostics, LLC, Rockville, MD.
 25. Quanterix Corporation. 2017. Simoa p24. Summary and explanation of the test. Quanterix Corporation, Lexington, MA.
 26. Bland JM, Altman DG. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 327:307–310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8).
 27. Manak MM, Hack HR, Nair SV, Worlock A, Malia JA, Peel SA, Jagodzinski LL. 2016. Evaluation of Hologic Aptima HIV-1 Quant Dx assay on the Panther system on HIV subtypes. *J Clin Microbiol* 54:2575–2581. <https://doi.org/10.1128/JCM.01350-16>.
 28. Chavez P, Wesolowski L, Patel P, Delaney K, Owen SM. 2011. Evaluation of the performance of the Abbott ARCHITECT HIV Ag/Ab combo assay. *J Clin Virol* 52(Suppl 1):S51–S55. <https://doi.org/10.1016/j.jcv.2011.09.010>.
 29. Nair SV, Kim HC, Fortunko J, Foote T, Peling T, Tran C, Nugent CT, Joo S, Kang Y, Wilkins B, Lednovich K, Worlock A. 2016. Aptima HIV-1 Quant Dx—a fully automated assay for both diagnosis and quantification of HIV-1. *J Clin Virol* 77:46–54. <https://doi.org/10.1016/j.jcv.2016.02.002>.
 30. Yek C, Massanella M, Peling T, Lednovich K, Nair SV, Worlock A, Vargas M, Gianella S, Ellis RJ, Strain MC, Busch MP, Nugent CT, Richman DD. 2017. Evaluation of the Aptima HIV-1 Quant Dx assay for HIV-1 RNA quantitation in different biological specimen types. *J Clin Microbiol* 55:2544–2553. <https://doi.org/10.1128/JCM.00425-17>.
 31. Alere North America. Alere Determine™ HIV-1/2 Ag/Ab combo. Alere North America, Orlando, FL.
 32. Stafylis C, Klausner JD. 2017. Evaluation of two 4th generation point-of-care assays for the detection of human immunodeficiency virus infection. *PLoS One* 12:e0183944. <https://doi.org/10.1371/journal.pone.0183944>.
 33. Livant E, Heaps A, Kelly C, Maharaj R, Samsunder N, Nhlangulela L, Karugaba P, Panchia R, Marrazzo J, Chirenje ZM, Parikh UM, VOICE Study Team. 2017. The fourth generation Alere™ HIV combo rapid test improves detection of acute infection in MTN-003 (VOICE) samples. *J Clin Virol* 94:15–21. <https://doi.org/10.1016/j.jcv.2017.06.006>.
 34. Cohen MS, Gay CL, Busch MP, Hecht FM. 2010. The detection of acute HIV infection. *J Infect Dis* 202(Suppl 2):S270–S277. <https://doi.org/10.1086/655651>.
 35. Duong YT, Mavengere Y, Patel H, Moore C, Manjengwa J, Sibandze D, Rasberry C, Mlambo C, Li Z, Emel L, Bock N, Moore J, Nkambule R, Justman J, Reed J, Bicego G, Ellenberger DL, Nkengasong JN, Parekh BS. 2014. Poor performance of the Determine HIV-1/2 Ag/Ab combo fourth-generation rapid test for detection of acute infections in a national household survey in Swaziland. *J Clin Microbiol* 52:3743–3748. <https://doi.org/10.1128/JCM.01989-14>.