

Cross-Platform Analysis of HIV-1 RNA Data Generated by a Multicenter Assay Validation Study with Wide Geographic Representation

Cheryl Jennings,^a Brian Harty,^b Suzanne Granger,^b Carrie Wager,^b John A. Crump,^c Susan A. Fiscus,^d and James W. Bremer^a

Rush Medical College, Department of Immunology and Microbiology, Chicago, Illinois, USA^a; New England Research Institutes, Inc., Watertown, Massachusetts, USA^b; Division of Infectious Diseases and International Health Department of Medicine, Duke University Medical Center, Durham, North Carolina, USA^c; and Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA^d

HIV-1 RNA quantitation continues to be extremely important for monitoring patients infected with HIV-1, and a number of assays have been utilized for this purpose. Differences in assay performance with respect to log₁₀ recovery and HIV-1 subtype specificity have been well documented for commercially available assays, although comparisons are usually limited to one or two assay platforms. Two new FDA-approved assays, the Roche Cobas AmpliPrep/Cobas TaqMan HIV-1 test (RT) and the Abbott RealTime HIV-1 assay (AR), that utilize real-time PCR have replaced previous HIV-1 RNA platforms. Inadequate detection of some strains of HIV-1 resulted in the addition of a new primer/probe set and the introduction of a second version of the RT assay. In this study, comparisons of assay performance between the different FDA-approved HIV-1 RNA assay platforms (both new and existing) were performed by using validation data that included both well-characterized virus stock and locally collected clinical samples. Laboratories across diverse geographical regions performed the validation testing and submitted data to the Virology Quality Assurance program (VQA) for analysis. Correlation values for clinical sample testing varied across the assay platforms ($r = 0.832$ to 0.986), and average log₁₀ recoveries for HIV-1 RNA controls (compared to the nominal value) ranged from -0.215 to 0.181 . These data demonstrate the need for use of one assay platform for longitudinal patient monitoring, but the data also reinforce the notion that no one assay is superior and that testing across platforms may be required for discordance reconciliation.

HIV-1 RNA quantitation (virus load testing) continues to be an important tool for monitoring HIV-1 infection (16, 29), and a number of different commercially available HIV-1 RNA assays are available for this purpose (1, 9, 10, 12, 20, 21, 26, 27, 28). HIV-1 RNA results are used to determine antiretroviral treatment efficacy, monitor safety and adherence, stratify patients enrolled in clinical trials, and predict virologic outcomes, such as mother-to-child transmission, treatment failure, and viral persistence in body compartments. In recent years, there has been a change in virus load methodologies, from endpoint PCR determinations to real-time PCR (10, 21, 27). The new real-time PCR methodologies offer a broad range of benefits over the traditional endpoint PCR, including higher precision, full automation with high throughput, expanded reportable range capabilities, and ever-decreasing lower limits of detection (20 to 40 copies/ml). However, with the improved sensitivity comes concern over an increased frequency of low-level viremia being detected (8, 18, 24) and the implications this may have for clinical management.

Four FDA-approved HIV-1 RNA assays (with multiple versions and platforms) were available in the United States from 2008 to 2011: the Abbott RealTime HIV-1 assay (AR); the Roche Cobas AmpliPrep/Cobas TaqMan HIV-1 test, versions 1 and 2 (RTv1 and RTv2); the Roche Amplicor HIV-1 Monitor test, version 1.5 (standard [STD] and ultrasensitive [US], the Cobas AmpliPrep [CAM], Cobas [COB], and microwell plate [MON] versions); and the Versant HIV-1 RNA 3.0 (bDNA) assay. An understanding of how the results generated from the different HIV-1 RNA assays compare is critical for all HIV-1 subtypes, especially since the genetic diversity of HIV-1 continues to evolve and underquantitation of HIV-1 continues to be documented (4, 11, 13, 23). Surveil-

lance and comparative studies worldwide (2, 17, 22, 25, 30, 31) are extremely important for documenting potential problems and forcing manufacturers to improve their assays in response to deficiencies (6, 7, 14). The high cost associated with the full automation of new real-time PCR assays has forced manufacturers to offer manual versions of the assays, but the impact of using manual extraction methods on precision across laboratories is not fully understood (5, 19). Many factors contribute to the variability of HIV-1 RNA measurements, including both systematic and biological variables (3). Parameters of assay performance that must be understood include accuracy, precision, sensitivity, specificity, and reproducibility across laboratories. More recent studies have shown good correlation among the current versions of the real-time PCR assays (6, 14, 29); the results of these studies demonstrate how comparative testing has forced assay improvement.

The Virology Quality Assurance program (VQA) is funded by the U.S. National Institutes of Health (NIH), National Institute of Allergy and Infectious Diseases (NIAID), Division of AIDS (DAIDS), to provide quality control, quality assurance, and quality assessment (proficiency testing) to laboratories performing HIV-1 virologic testing for NIH-sponsored clinical trials (32). The

Received 2 March 2012 Returned for modification 17 May 2012

Accepted 5 June 2012

Published ahead of print 12 June 2012

Address correspondence to Cheryl Jennings, cheryl_jennings@rush.edu.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JCM.00578-12

VQA developed an HIV-1 RNA assay validation plan that was utilized by client laboratories to validate newly installed FDA-approved HIV-1 RNA assays. Data from these validation studies were combined to summarize the overall performance of the new real-time PCR assays (RTv1, RTv2, and AR) with the previous endpoint PCR assays (STDCAM, USCAM, STDCOB, USCOB, STDMON, and USMON). This evaluation is novel because it combines the testing of a well-characterized control stock across multiple assay platforms with the testing of locally collected clinical samples from laboratories representing diverse geographical regions.

(The VQA validation plan was presented, in part, as a poster at the 15th Conference on Retroviruses and Opportunistic Infections, Montreal, CA, February 2009, poster 1001.)

MATERIALS AND METHODS

The validation plan required each laboratory to run VQA HIV-1 RNA copy controls with known nominal values, which were based on the dilution of a well-characterized stock virus (32) into commercially available pooled human plasma (sodium citrate treated) or pooled human serum to which a liquid EDTA solution (500 mM; Sigma) was added (final concentration, 6 mM). The controls used in the validation testing were first produced in human plasma (citrate) and then in serum plus 6 mM EDTA to accommodate any assay (specifically, RTv2) that requires the use of EDTA plasma for sample collection. Titration experiments conducted by the VQA demonstrated that the \log_{10} recoveries in samples produced with serum containing 6 mM EDTA were similar to those noted in samples created using EDTA-treated plasma (data not shown). Plasma was derived from locally obtained donor samples collected in EDTA Vacutainers. VQA controls were combined with clinical samples according to predefined templates that were tested over five separate days. Each template was tested using the new and existing comparator assays. All assays were performed according to the package insert for each assay that was used. Only valid data were included in these analyses. Sample validity was defined by each assay manufacturer and included requirements for kit control performance to fall within the specified range, internal control performance, and quantitation standard performance. A quantitative result was required for all samples, except for samples with low copy numbers. Samples with a titer above the upper limit of detection for the assay were to be diluted, and HIV-1 RNA values were calculated from retesting of the diluted sample.

Statistics. Data generated from VQA HIV-1 RNA controls were used for analyses of accuracy, precision, linearity, and sensitivity, and locally collected clinical samples were used for analyses of specificity and clinical sample correlations.

For each control sample, the \log_{10} recovery was computed as the \log_{10} -transformed result minus the \log_{10} -transformed nominal value. Linear mixed effects models were fit to the \log_{10} recovery for each sample from each assay and control matrix and included precision components for intrarun, interrune, and interlaboratory variations as well as fixed effects corresponding to concentration-specific estimates of recovery (CSERs). Intralaboratory total assay standard deviations (SD) were computed as the square root of the sum of the intrarun and interrune variance. A critical threshold of the intralaboratory total assay SD was set at 0.15, which corresponded to the clinical target of being able to detect a 5-fold change in virus load with 90% power and a type I error rate of 95%. Overall accuracy was estimated using a design-weighted average of the CSERs.

The three linearity parameters assessed in the individual validation plan included the slope and residual standard deviation (SD_{resid}) of a linear fit to the CSERs as well as the standard errors of the means (SEM) of CSERs. Acceptance bounds for these linearity parameters were derived from data simulations and pertained to the maximum variation that would be expected for an individual validation (34 control samples from one laboratory across 5 runs) having a total assay SD lower than the target

of 0.15. In individual validations, linearity problems were noted if all three parameters exceeded their respective acceptance bounds. The mean estimates of the linearity parameters across all laboratories were estimated from the CSERs derived from the linear mixed effects model used for the precision analysis (described above). That model also included laboratory-specific estimates of linear deviations from the CSERs as part of the interlaboratory SD component, and these were used to derive estimates of the interlaboratory SD for the slope and SEM parameters.

Sensitivity evaluations were performed by calculating the percentage of samples that yielded detectable results for samples with a nominal value near the limit of detection for each assay. The result needed only to be a qualitative result, indicating that HIV-1 RNA was detected. Samples with nominal values of 25 and 50 copies/ml were chosen for these studies, even though the lower limits of detection for each assay platform do vary. The main purpose of these analyses was to evaluate performance based on the respective kit manufacturer claims; the secondary purpose was to evaluate performance across the control matrices.

The results generated for seropositive clinical samples were compared across each validated pair of platforms. Discordance was defined as a result that differed between the two assays with respect to qualitative and quantitative detection. For the samples that had quantitative results for both platforms, the correlation was assessed, and the difference between \log_{10} -transformed results for the two assay platforms was computed. Correlations below 0.95 and individual differences that exceeded 5-fold ($0.7 \log_{10}$) were deemed problematic. Bland-Altman plots were constructed, and a smooth trend along with a 95% confidence band that controlled for lab-specific variation was fit to the differences across the mean concentration. In addition, the 90% limits of agreement (LOA) were included on these plots. The 90% LOA is a measure of the overall degree of agreement between the two assays and can be interpreted as the range in which we expect 90% of the differences between the 2 assays to fall. It is calculated as the average difference across all seropositive samples ± 1.96 times the standard deviation of the differences. The expectation is that this interval will fall within the $\pm 0.7 \log_{10}$ range.

RESULTS

Assays and data sets. The nominal value and number of replicates of VQA HIV-1 RNA controls and unique clinical donor samples tested for each validation evaluation were predefined. These included replicates of HIV-1 RNA controls with nominal values for copies/ml (with n indicated in parentheses) of 25 (4), 50 (8), 1,500 (9), 15,000 (9), 150,000 (9), and 1,500,000 (7), as well as locally collected plasma from 20 unique HIV-1-uninfected and 39 HIV-1-infected subjects. All samples were tested using the same templates and were assayed on the new and comparator assay platforms over five consecutive days. The number of validation data sets received for each assay and control matrix is provided in Table 1. A total of 40 unique data sets were received from 32 laboratories, representing 15 countries enrolled in the VQA HIV-1 RNA proficiency testing program (Table 2) over 3 years (2008 to 2011). Thirteen data sets were used to validate the installation of the AR assay using the automated m2000sp sample processing instrument (ARAUTO); 11 data sets were submitted for the validation of the AR assay using the manual extraction procedure (ARMAN); 12 data sets were submitted for the validation of the RTv1 assay; and 7 data sets were submitted for the validation of the RTv2 assay. Comparator assays varied by laboratory and included a range of FDA-approved assays. For five data sets, comparator testing was performed by the VQA or other contracted laboratory by using a previously validated, FDA-approved assay; all other data sets included data generated by one laboratory for both the new and comparator assay testing.

Accuracy. Accuracy for each assay and control matrix, as esti-

TABLE 1 Validation panels tested per assay and control matrix

Assay	Abbreviation	No. of validation panels tested (by control matrix)		
		Citrate	EDTA	Total
Abbott RealTime 0.6 ml (automated extraction)	ARAUTO	10	3	13
Abbott RealTime 0.6 ml (manual extraction)	ARMAN	3	8	11
Roche TaqMan v1.0	RTv1	9	3	12
Roche TaqMan v2.0	RTv2	1	6	7
Roche standard Cobas-Ampliprep	STDCAM	5	0	5
Roche standard Cobas	STDCOB	2	4	6
Roche standard MWP	STDMON	0	5	5
Roche standard (all platforms combined)	STD	7	9	16
Roche Ultrasensitive Cobas-Ampliprep	USCAM	3	2	5
Roche Ultrasensitive Cobas	USCOB	1	1	2
Roche Ultrasensitive MWP	USMON	4	2	6
Roche Ultrasensitive (all platforms combined)	US	8	5	13

mated by the average \log_{10} recovery across the four highest nominal concentrations of VQA HIV-1 RNA copy number controls, is presented in Table 3. A \log_{10} recovery of zero suggested that there was no difference between the result obtained with a given assay and the nominal value of the control. A negative \log_{10} recovery suggested that the assay reported values lower than the nominal value, and a positive \log_{10} recovery indicated the result was higher than the nominal value. Average \log_{10} recovery varied across all assays. Average \log_{10} recovery for citrate controls ranged from -0.124 for USMON to 0.431 for RTv2. Average \log_{10} recovery for EDTA controls ranged from -0.215 for ARAUTO to 0.181 for USCAM. In some cases, differences in average \log_{10} recoveries were noted between platforms of the same assay; for EDTA controls, average \log_{10} recoveries for USMON, USCOB, and USCAM were -0.119 , 0.107 , and 0.181 , respectively. Differences in average \log_{10} recoveries were also noted between extraction methods of the same assay; the average \log_{10} recoveries for EDTA controls were -0.215 and -0.021 for the ARAUTO and ARMAN, respectively.

Precision. Intralaboratory total SD of \log_{10} recoveries were higher than the target of 0.15 for the combined data sets generated using the STDCAM assay for controls produced in citrate and in the combined data sets generated with USCAM for controls produced in EDTA (Table 3). The combined data for all platforms of the STD Roche Amplicor HIV-1 Monitor test in citrate and for all platforms of the US Roche Amplicor HIV-1 Monitor test in EDTA showed an elevated intralaboratory total SD. Intrarun variation was the main cause for this elevated precision statistic regardless of the matrix. An interlaboratory component of variation was also evaluated for this analysis. Adding the interlaboratory variation to the inter- and intrarun variations provided an overall total assay standard deviation that was estimated for each assay. The overall total assay standard deviations ranged from 0.083 to 0.266 for citrated controls and from 0.159 to 0.225 for EDTA-treated con-

TABLE 2 Participating laboratories

Data set no.	City and state or country	Date (mo/day/yr)	New assay	Comparator assay ^a
1	Chapel Hill, NC	11/20/08	ARAUTO	USMON
2	Chicago, IL	11/20/08	ARAUTO	USMON
3	Chicago, IL	11/20/08	ARAUTO	USCOB
4	Chicago, IL	09/01/10	RTv2	USCOB
5	Chiang Mai, Thailand	11/20/08	ARAUTO	USMON
6	Durban, South Africa	11/20/08	RTv1	STDCAM
7	Lusaka, Zambia	11/20/08	RTv1	STDCAM
8	Lusaka, Zambia	04/20/10	ARAUTO	STDCAM
9	Lusaka, Zambia	05/27/10	ARAUTO	RTv1
10	Moshi, Tanzania	11/20/08	ARAUTO	USCOB*
11	Moshi, Tanzania	11/20/08	ARMAN	
12	Durham, NC	11/20/08	RTv1	USMON*
13	Baltimore, MD	04/21/09	ARAUTO	USMON
14	Johannesburg, South Africa	08/21/09	RTv1	STDCAM
15	Johannesburg, South Africa	07/05/11	RTv2	RTv1
16	Johannesburg, South Africa	07/05/11	ARAUTO	RTv1
17	Kingston, Jamaica	09/23/09	RTv1	USCAM
18	Johannesburg, South Africa	10/16/09	ARAUTO	STDCAM
19	Johannesburg, South Africa	10/16/09	RTv1	
20	Johannesburg, South Africa	10/16/09	RTv2	
21	Bangkok, Thailand	02/19/10	RTv1	USCAM
22	Chiang Mai, Thailand	03/9/10	ARMAN	STDCOB
23	Chennai, India	03/15/10	ARMAN	STDCOB
24	Khonkaen, Thailand	04/20/10	RTv1	USCAM*
25	Gaborone, Botswana	05/27/10	ARAUTO	USCAM
26	Pittsburgh, PA	07/09/10	RTv2	USMON
27	Johannesburg, South Africa	07/28/10	ARAUTO	USCAM
28	Eldoret, Kenya	07/28/10	ARMAN	STDMON
29	Harare, Zimbabwe	08/24/10	ARMAN	STDMON
30	Chicago, IL	09/01/10	RTv2	RTv1*
31	Buenos Aires, Argentina	11/19/10	RTv2	USCAM
32	Nairobi, Kenya	11/19/10	ARMAN	STDMON
33	Pune, India	12/06/10	ARMAN	STDCOB
34	Entebbe, Uganda	12/06/10	RTv2	STDCOB
35	Port au Prince, Haiti	01/31/11	ARMAN	STDCOB
36	Birmingham, AL	01/31/11	RTv2	RTv1
37	Pune, India	03/14/11	ARMAN	STDCOB
38	Rio de Janeiro, Brazil	03/14/11	ARMAN	USMON*
39	Blantyre, Malawi	04/14/11	ARAUTO	STDMON
40	Kampala, Uganda	05/17/11	ARMAN	STDMON

^a *, in five comparator data sets, comparator testing was performed by the VQA or other contracted laboratory using a previously validated, FDA-approved assay; all other data sets included data generated by one laboratory for both the new and comparator assay testing.

trols. A definitive target for this precision parameter has not yet been established.

Linearity. The linearity parameters (slope, $SD_{residual}$, and SEM) are presented in Table 4. All parameters should be close to zero when there is constant recovery across concentrations. For the RTv2 assay in citrate, the mean slope and SEM parameters exceeded the target value; however, it should be noted that the assessment was based upon data from one laboratory with a restricted range of nominal concentrations. For the RTv2 assay in EDTA, the large negative slope across concentrations indicates a decreasing linear bias across concentrations. The other linearity parameters were within range for all platforms, indicating no ev-

TABLE 3 Accuracy and precision statistics by control matrix and assay

Control matrix and assay	Accuracy (avg log ₁₀ recovery)	Precision (SD) ^a				Overall total
		Intrarun	Interrun	Intralaboratory total ^b	Interlaboratory	
Citrate						
ARAUTO	-0.046	0.059	0.052	0.078	0.061	0.099
ARMAN	0.012	0.042	0.054	0.069	0.045	0.083
RTv1	0.309	0.111	0.016	0.112	0.059	0.127
RTv2	0.431	0.065	0.041	0.076	—	—
STDCAM	0.090	0.172	0.022	0.173	0.026	0.175
STDCOB	0.044	0.131	0.000	0.131	0.067	0.147
STDMON ^d						
STD ^c	0.078	0.161	0.024	0.163	0.038	0.168
USCAM	0.263	0.115	0.056	0.127	0.168	0.211
USCOB	-0.011	0.093	0.076	0.120	—	—
USMON	-0.124	0.146	0.026	0.148	0.133	0.200
US ^c	0.037	0.132	0.052	0.142	0.225	0.266
EDTA						
ARAUTO	-0.215	0.051	0.036	0.062	0.184	0.194
ARMAN	-0.021	0.067	0.066	0.094	0.130	0.160
RTv1	0.040	0.078	0.043	0.089	0.207	0.225
RTv2	0.133	0.072	0.045	0.085	0.134	0.159
STDCAM ^d						
STDCOB	0.075	0.121	0.013	0.122	0.165	0.205
STDMON	0.081	0.111	0.056	0.124	0.141	0.188
STD ^c	0.080	0.115	0.047	0.124	0.139	0.187
USCAM	0.181	0.140	0.071	0.157	0.041	0.162
USCOB	0.107	0.087	0.098	0.131	—	—
USMON	-0.119	0.141	0.000	0.141	—	—
US ^c	0.056	0.136	0.066	0.151	0.151	0.214

^a Boldface values indicate when the target threshold was exceeded. —, statistics could not be computed because data were received from only one laboratory for this assay.

^b Each laboratory was included in the precision component model, but the interlaboratory precision component was not included in the calculation of the total assay standard deviation.

^c Data sets STD and US represent the combined standard or ultrasensitive data across all platforms for Roche Amplicor HIV-1 Monitor testing, respectively.

^d Data were not received for STDMON in citrate or for STDCAM in EDTA.

idence of either linear or nonlinear deviation from a flat trend across concentrations. Figure 1 provides a depiction of the linearity assessment for all laboratories for each assay and control matrix. Overall linear trends were more similar across assays for EDTA controls than for citrated controls.

Assay sensitivity. The lower limits of detection were evaluated across assays using VQA controls with nominal values of 25 and 50 copies/ml (Table 5). Detection was based on the corresponding manufacturer definition. Briefly, a sample was detected if a minimum optical density (CAM, COB, and MON) or if a minimum cycle threshold (RTv1, RTv2, and AR) was obtained. The ultrasensitive endpoint PCR assays (USCAM, USCOB, and USMON) have a lower detection limit of 50 copies/ml; the standard endpoint PCR assays (STDCAM, STDCOB, and STDMON) have a lower detection limit of 400 copies/ml; the real-time PCR assays have lower detection limits of 20, 40, and 48 copies/ml for RTv2, AR (MAN and AUTO), and RTv1, respectively. Detection rates for the standard endpoint PCR assays were very low (0 to 31%) for both control concentrations, regardless of matrix, but these controls had nominal values that were well below the lower detection limit of the assays.

Excluding the standard endpoint PCR assays, a total of 96 to 100% of the citrated control samples with nominal values of 50 copies/ml were detected, compared to only 63 to 100% of the

EDTA controls. The ARAUTO, RTv1, and USCAM assays detected 71, 83, and 63% of EDTA controls with a nominal value of 50 copies/ml, respectively; these were all significantly lower than would be expected based on a 95% detection rate at 50 copies/ml ($P < 0.001$, $P = 0.03$, and $P < 0.001$, respectively). Samples with a nominal value of 25 copies/ml were detected in 81 to 100% and 38 to 100% of the citrated and EDTA controls, respectively (Table 5).

For the endpoint PCR assays, one false-positive result each was noted for the USCAM assay (63 copies/ml) and for the USMON assay (detected, <50 copies/ml) on specimens that were confirmed negative by HIV-1 antibody testing performed by the testing laboratory, using an FDA-approved assay. No other false-positive results were noted with any other assay platform. Carryover effects, e.g., from running a sample with no HIV-1 RNA immediately following a sample with a very high virus load, were assessed by running several negative samples immediately following samples with a nominal value of 1,500,000 copies/ml. No carryover effect was noted in any of the HIV-1 RNA assays (data not shown).

Clinical correlations. Table 6 summarizes the comparisons from the seropositive clinical sample testing. The percentage of discordant results between the new and comparator assays ranged from 1 to 10%, 1 to 8%, 2 to 5%, and 1 to 9% for ARAUTO, ARMAN, RTv1, and RTv2, respectively, for samples near the lower limits of detection for the assays. There were four results

TABLE 4 Linearity statistics by control matrix and assay^a

Control matrix and assay	Slope			SEM	
	Mean	SD	SD _{resid}	Mean	SD
Citrate					
ARAUTO	0.014	0.019	0.025	0.019	0.025
ARMAN	0.002	0.016	0.016	0.003	0.020
RTv1	0.027	0.018	0.051	0.035	0.023
RTv2	-0.096	0.035	0.030	0.124	0.045
STDCAM	0.028	0.042	0.073	0.036	0.055
STDCOB	-0.014	0.041	0.048	0.017	0.053
STDMON ^b					
STD	0.039	0.037	0.018	0.051	0.048
USCAM	-0.036	0.016	0.049	0.046	0.020
USCOB	0.031	0.032	0.065	0.040	0.041
USMON	-0.020	0.032	0.010	0.025	0.041
US	0.014	0.019	0.025	0.019	0.025
EDTA					
ARAUTO	0.024	0.014	0.024	0.031	0.017
ARMAN	-0.011	0.037	0.016	0.014	0.048
RTv1	0.008	0.023	0.035	0.011	0.030
RTv2	-0.065	0.015	0.062	0.083	0.019
STDCAM ^b					
STDCOB	0.009	0.045	0.062	0.012	0.058
STDMON	0.005	0.035	0.073	0.007	0.045
STD	-0.032	0.056	0.020	0.042	0.073
USCAM	-0.010	0.056	0.046	0.013	0.072
USCOB	-0.015	0.054	0.034	0.020	0.069
USMON	0.014	0.042	0.021	0.018	0.054
US	0.024	0.014	0.024	0.031	0.017

^a The targets used for the individual validations were 0.056 for the absolute value of the slope, 0.096 for SD_{resid}, and 0.091 for SEM. Boldface values indicate when the target threshold was exceeded by the mean estimate for the indicated parameter.

^b Data were not received for STDMON in citrate or for STDCAM in EDTA.

where the virus load in a sample was missed (qualitatively detected or not detected) with one assay and the virus load reported with the other assay was more than 0.7 log₁₀ RNA copies/ml (5-fold) higher than the lower detection limit for that assay. In one example, the result for a sample that was reported as 550 copies/ml with the USCAM assay yielded a result of <40 copies/ml HIV-1 RNA detected with the ARAUTO assay. In two other comparisons, the ARAUTO assay reported 321 and 160 copies/ml for samples that were undetectable or qualitatively detected (<50 copies/ml) with the USMON assay. In the fourth example, a result of 1,870 copies/ml was reported with RTv2 and a result of <400 copies/ml HIV-1 RNA detected was reported with STDCAM. None of the other qualitatively discordant results deviated by more than 0.7 log₁₀ copies/ml from the expectation of the assay's respective lower detection limit. A total of 413, 366, 261, and 278 clinical sample pairs that yielded quantitative results for both assays were tested on ARAUTO, ARMAN, RTv1, and RTv2, respectively. From those totals, 29 (7%), 11 (3%), 13 (5%), and 7 (3%) of the samples yielded results that differed by more than 0.7 log₁₀ RNA copies/ml between the two assays.

The new and comparator assays showed a range of correlation values across all comparisons made (Fig. 2). All assay comparisons that included versions of the standard Roche Amplicor HIV-1 Monitor assay (STDCAM, STDCOB, and STDMON) or the ultra-sensitive Roche Amplicor HIV-1 Monitor assay (USCAM, USCOB, and USMON) as a comparator were each, respectively, combined into one graph, because most clinicians do not differentiate between the different extraction methods of these assays. Correlation values (*r*) ranged from 0.832 to 0.949 for comparisons involving ARAUTO, from 0.909 to 0.950 for comparisons submitted for ARMAN, from 0.882 to 0.962 for comparisons performed against RTv1, and from 0.921 to 0.986 for comparisons involving RTv2. All of these correlation values were highly significant (*P* < 0.0001). Bland-Altman plots for the different assay combinations are shown in Fig. 3. Individual differences that exceeded 0.7 log₁₀

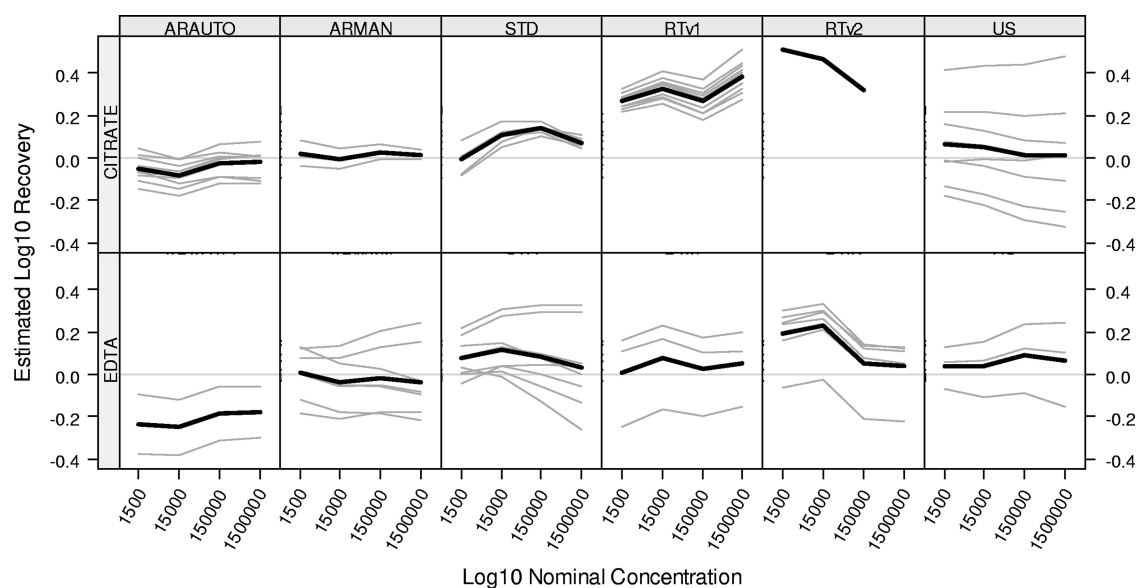


FIG 1 Linearity assessment summary by kit. CSERs were derived from the models described in Table 4. The model allows for a nonlinear trend across concentration that is common to all labs for a given kit (thick black lines). Individual labs may have linear deviations from that overall nonlinear trend (gray lines). A horizontal reference line is displayed at zero.

TABLE 5 Qualitatively detectable results per manufacturer definition,^a by control matrix

Control matrix and assay	Results for samples with nominal concn of:				P value ^b
	25 copies/ml		50 copies/ml		
	No. of samples	% qualitatively detectable	No. of samples	% qualitatively detectable	
Citrate					
ARAUTO	40	90	80	100	1.00
ARMAN	12	92	24	96	0.71
RTv1	36	92	72	96	0.70
RTv2	4	100	8	100	1.00
STDCAM	20	25	40	28	<0.001
STDCOB	8	0	16	13	<0.001
STDMON ^c					
STD	28	18	56	23	<0.001
USCAM	12	83	24	96	0.71
USCOB	4	100	8	100	1.00
USMON	16	81	32	97	0.81
US	32	84	64	97	0.84
EDTA					
ARAUTO	12	67	24	71	<0.001
ARMAN	32	88	64	100	1.00
RTv1	12	67	24	83	0.03
RTv2	24	88	48	98	0.91
STDCAM ^c					
STDCOB	16	31	32	31	<0.001
STDMON	20	0	32	20	<0.001
STD	36	14	72	25	<0.001
USCAM	8	38	16	63	<0.001
USCOB	4	100	8	100	1.00
USMON	4	100	8	100	1.00
US	16	69	32	81	<0.01

^a The validation controls used for this testing were designed to evaluate assays with ultralow detection limits (50 copies/ml or less). Standard Roche assays (STDMON) have lower detection limits of 400 copies/ml. A blank indicates no data were received for this assay.

^b A P value is provided for each assay to indicate if the detection rate of the sample with a nominal value near the lower detection limit of the assay (e.g., 50 copies/ml) was significantly different (too many false negatives) than what would be expected by chance alone for a 95% detection rate.

^c Data were not received for STDMON in citrate or for STDCAM in EDTA.

RNA copies/ml were deemed problematic. For two of the comparisons (STD versus RTv2 and US versus RTv2), the average difference between paired results trended significantly away from zero for a portion of the average concentration range. The 90% limits of agreement exceeded the $\pm 0.7 \log_{10}$ limit for four of the comparisons (STD versus ARAUTO, US versus ARMAN, STD versus RTv1, and STD versus RTv2), indicating a lack of evidence that the two platforms are equivalent for these comparisons.

DISCUSSION

Validation testing must be done whenever a new test is introduced, in order to verify the assay specifications set forth by the manufacturer. A good assay validation plan must include the use of a well-characterized control material for purposes of verifying assay specifications as well as use of locally collected clinical samples for purposes of clinical correlation studies. A number of investigators have reported significantly discordant virus load results in samples containing non-B HIV-1 subtypes across the new

HIV-1 RNA real-time PCR assays (4, 11, 13, 23, 26, 31). Such correlation studies are extremely important and actually resulted in the modification from the RTv1 to the RTv2 assay (6, 7, 12). For correlation studies, samples must be tested using the new assay platform, and those results must be confirmed against a previously validated assay. This is important from a clinical perspective to confirm if rebaselining of patient results will be necessary. For this analysis, validation data were combined to provide novel information on HIV RNA assay performance across a variety of assays, including multiple platforms within a given assay. Despite the facts that the number of validation data sets varied between new and comparator assay platforms and there was no confirmation of the HIV subtype in the clinical sample subsets, these data provide a rich database generated by a diverse group of testing laboratories. Historically, HIV-1 RNA assays did not mandate the use of EDTA as a specimen anticoagulant; however, the RTv1 and RTv2 assays now require the use of EDTA for specimen collection. Preliminary analyses based on proficiency testing data suggest that the \log_{10} recovery of data generated on the RTv1 assay is significantly higher than that noted in other assays; this increased \log_{10} recovery was later attributed to the use of citrated plasma for panel productions. The data generated for validation confirmed this observation (Table 3). In general, the average \log_{10} recovery was higher in citrated controls than in EDTA controls across all real-time PCR assays. One hypothesis is that EDTA is carried through the extraction and chelates the manganese in the PCR master mix, changing the efficiency of the PCR. The problem appears to be greater in assays that utilize magnetic iron-containing particles for extraction. Iron has a strong affinity for EDTA, so it may help to carry the EDTA through the extraction, which can affect PCR efficiency. Both the AR and RT assays utilize extraction reagents that contain iron (AR uses iron oxide and RT uses magnetic glass particles). The magnitude of the effect on \log_{10} recovery varies across the platforms and supports the notion that sample dilutions must be made using diluents that contain EDTA. However, the effect on accuracy noted in VQA controls produced in citrated plasma did not translate to clinical samples derived from citrated whole blood collected in a Vacutainer, which were diluted (15%) by the liquid anticoagulant (acid citrate dextrose or cell preparation tube anticoagulant; Becton, Dickinson). The dilution apparently negated the anticoagulant effect (data not shown).

Many components contribute to the variability noted in a real-time virus load measurement, including biological variability, systematic variability, and assay variability (3). When monitoring patients infected with HIV-1 over time, it is important to understand the magnitude of the components of variation and to know what change in virus load is clinically relevant. Results obtained on two different HIV-1 RNA platforms may disagree due to differences in technologies and varying degrees of genetic variability of HIV-1, which cause problems with primer and probe hybridization in these molecular-based HIV-1 RNA assays (6, 7, 23, 30). In previously published guidance documents, changes in HIV-1 RNA of $<0.5 \log_{10}$ HIV-1 RNA copies/ml (3-fold) were not considered clinically relevant (29). This criterion was based on the assumption that the two measurements were generated with the same assay. In the evaluation described here and in our routine validation testing, we consider differences of $>0.7 \log_{10}$ HIV-1 RNA copies/ml (5-fold) to be clinically relevant differences. The larger value was applied because it contains the added variation component associated with using two different assay platforms for

TABLE 6 HIV-1 RNA estimate summary for HIV-1-seropositive clinical sample pairs by assay combination

New assay	Comparator assay	No. (%) of HIV-1 RNA copies/ml, according to clinical sample pairs ^a						No. (%) of discordant pairs ^c
		Both TND	1 TND, 1 Qual ^b	1 TND, 1 Quant ^b	1 Qual, 1 Quant ^b	Both Qual	Both Quant	
ARAUTO	RTv1	7 (9)	2 (3)		1 (1)	2 (3)	66 (85)	1 (2)
ARAUTO	STDCAM	11 (14)	8 (10)	5 (6)	6 (8)		48 (62)	4 (8)
ARAUTO	STDMON						39 (100)	9 (23)
ARAUTO	STD	11 (9)	8 (7)	5 (4)	6 (5)		87 (74)	13 (15)
ARAUTO	USCAM	1 (1)			2 (3)		75 (96)	7 (9)
ARAUTO	USCOB	2 (3)	2 (3)		2 (3)	2 (3)	70 (90)	2 (3)
ARAUTO	USMON	15 (10)	12 (8)	3 (2)	6 (4)	5 (3)	115 (74)	6 (5)
ARAUTO	US	18 (6)	14 (4)	3 (1)	10 (3)	7 (2)	260 (83)	15 (6)
Total for ARAUTO comparisons		36 (7)	24 (5)	8 (2)	17 (3)	9 (2)	413 (81)	29 (7)
ARMAN	STDCOB	9 (5)	2 (1)		7 (4)		165 (90)	7 (4)
ARMAN	STDMON	8 (5)	2 (1)	12 (8)	2 (1)		136 (85)	2 (1)
ARMAN	STD	17 (5)	4 (1)	12 (3)	9 (3)		301 (88)	9 (3)
ARMAN	USCOB	1 (3)	1 (3)		1 (3)	1 (3)	35 (90)	1 (3)
ARMAN	USMON	5 (13)	1 (3)	2 (5)			30 (79)	1 (3)
ARMAN	US	6 (8)	2 (3)	2 (3)	1 (1)	1 (1)	65 (84)	2 (3)
Total for ARMAN comparisons		23 (5)	6 (1)	14 (3)	10 (2)	1 (0.2)	366 (87)	11 (3)
RTv1	STDCAM	16 (10)	3 (2)	7 (4)	7 (4)		123 (79)	11 (9)
RTv1	STD	16 (10)	3 (2)	7 (4)	7 (4)		123 (79)	11 (9)
RTv1	USCAM	1 (1)	2 (2)			1 (1)	113 (97)	1 (1)
RTv1	USMON	10 (26)	2 (5)		1 (3)	1 (3)	25 (64)	1 (4)
RTv1	US	11 (7)	4 (3)		1 (1)	2 (1)	138 (88)	2 (1)
Total for RTv1 comparisons		27 (9)	7 (2)	7 (2)	8 (3)	2 (1)	261 (84)	13 (5)
RTv2	RTv1		3 (3)		7 (6)	1 (1)	106 (91)	1 (1)
RTv2	STDCAM	7 (18)		8 (21)	2 (5)		22 (56)	3 (14)
RTv2	STDCOB	6 (15)	2 (5)				31 (79)	1 (3)
RTv2	STD	13 (17)	2 (3)	8 (10)	2 (3)		53 (68)	4 (8)
RTv2	USCAM						39 (100)	2 (5)
RTv2	USCOB		1 (3)		2 (5)		36 (92)	0 (0)
RTv2	USMON	59 (42)	13 (9)	20 (14)	2 (1)	1 (1)	44 (32)	0 (0)
RTv2	US	59 (27)	14 (6)	20 (9)	4 (2)	1 (0)	119 (55)	2 (2)
Total for RTv2 comparisons		72 (17)	19 (5)	28 (7)	13 (3)	2 (0.5)	278 (67)	7 (3)

^a TND, target not detected; Qual, qualitative; Quant, quantitative.

^b Results in these categories represent discordant results for samples with low virus titers.

^c The results in this category reflect discordance in samples that yielded quantitative results on both the new and comparator virus load assay for which the log₁₀ recovery rate difference was >0.7. The log₁₀-transformed results from the comparator assay were subtracted from the log₁₀-transformed results obtained with the new assay. Blank cells in the table indicate that there were no clinical sample pairs that fit the category description.

generating a virus load. For clinical trial testing, many laboratories from diverse geographical locations contribute data, and interlaboratory variation is yet another component of variation that must be taken into consideration. In this evaluation, the new real-time PCR assays offered better intralaboratory precision than did the endpoint PCR assays. When the interlaboratory component was added to the precision statistic for real-time PCR assays, total assay precision was increased for all assays, but a target for this precision statistic is yet to be defined. Higher interlaboratory differences were noted in controls produced in EDTA, but this may have been due to differences in sample sizes. The additional variation noted between laboratories could be attributed to kit lot variability, equipment variability, or inexperience with performing testing. Real-time monitoring of assay performance through

the use of external quality control materials and interlaboratory performance tracking through external proficiency testing programs are very important to help monitor this variation.

The linearity analyses presented in this evaluation provide a novel, quantifiable measure for the assessment of assay linearity. There are three distinct concepts of linearity that are routinely assessed in assay validation: the linearity of a diluted sample, performance (accuracy and precision) within a linear range, and constant recovery across a range of concentrations. The first concept is not clinically relevant in HIV-1 RNA testing, especially now that real-time PCR assays have such large linear ranges (20 to 10,000,000 copies/ml). The second concept is already addressed with the assessments of precision and accuracy of well-characterized controls that span the linear range of the assays. The third

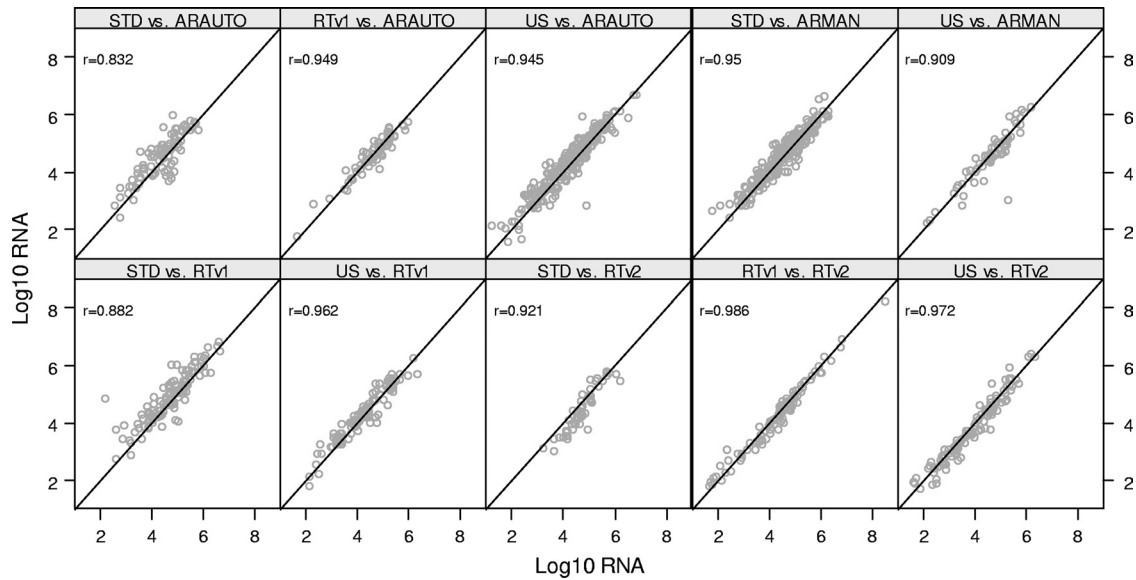


FIG 2 Clinical sample regression plots. Correlations are shown between the comparator assay (y axis) and the new assay (x axis) for each pair of the comparisons. Multiple labs are represented within each panel, and the correlations were computed after adjusting for within-lab deviations.

concept, constant recovery, was assessed in this evaluation by including the slope, SD_{resid} , and SEM parameters presented in Table 4; the RTv2 assay demonstrated an elevated parameter for slope and SEM in citrate and an elevated slope parameter in EDTA.

The RTv1 assay tended to run slightly lower at the low end of the linear range, giving it a positive slope from low to high concentrations; the RTv2 assay demonstrated a negative slope with

higher recovery at the low end. This may be a consequence of trying to improve the lower detection limit for this assay. The ARAUTO and ARMAN assays tended to have constant recoveries across the linear range, but there was more variability in linearity parameters between laboratories with the ARMAN assay, as depicted by the intersecting lines in the linearity plots. The implication is that variability in assay calibration may be due to variability

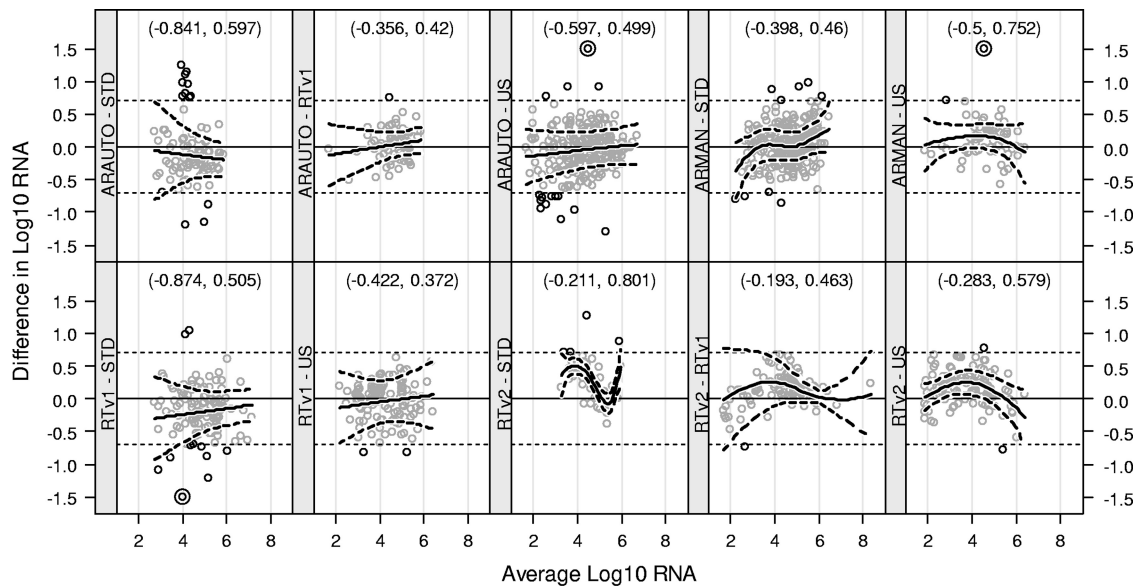


FIG 3 Bland-Altman plots. Within each panel, the difference between measurements for each sample (y axis) was computed as the value from the new assay minus that for the old assay. Horizontal dotted reference lines at ± 0.7 targets flag individual differences that exceeded 5-fold. The scale was truncated at ± 1.5 to allow greater resolution of the trend, and three measurements having values outside that range are displayed using double concentric circles (for US ARAUTO, the actual value is 2.257; for US ARMAN, the actual value is 2.385; for RTv1 STD, the actual value is -2.823). Nonlinear trends (thick solid lines) were fit to the data within each panel, along with an interval within which trend lines for 90% of the labs were expected to lie (thick dashed lines). Intervals which exceeded ± 0.7 for some portion of the concentration range indicate that the average trend for 5% of the labs may exceed 5-fold for those concentrations. Intervals that exclude zero for some portion of the concentration range indicate that the average trend differed significantly ($\alpha = 0.10$) from zero, but perhaps the trend was not clinically relevant. A 90% LOA (based upon the mean difference \pm the intralaboratory SD for each comparison) is provided at the top of each panel.

associated with the manual extraction. Since assay performance characteristics with respect to accuracy, precision, and linearity are all interrelated, major failures in assay performance are typically associated with failures across multiple criteria. Therefore, elevations in any one parameter, such as the slope or SEM, will most likely have little effect on clinical management, especially if other criteria, such as assay precision, fall well below expected targets. While there were a couple of laboratories that manifested multiple problems in their validation testing, most problems were fixed with retraining and retesting, suggesting the problems were associated with technical performance, not necessarily assay performance. Since EDTA-anticoagulated Vacutainer tubes are typically used for collection of whole blood for HIV-1 viral loads, the data generated using EDTA controls are more relevant to \log_{10} recovery data for the assays. The data in this analysis showed average \log_{10} recoveries of -0.215 , -0.021 , 0.040 , and 0.133 for the ARAUTO, ARMAN, RTv1, and RTv2 assays, respectively (Table 3). This is consistent with the findings reported by van Rensburg et al., who showed that the AR assay underreported the WHO standard by 0.2 to $0.3 \log_{10}$ RNA copies/ml and the RTv2 assay overreported the WHO standard by 0.17 to $0.31 \log_{10}$ RNA copies/ml (30). This further demonstrates the importance of using one assay platform for monitoring longitudinal data for a given subject, as well as the importance of performing assay validations to determine the impact of switching virus load platforms.

The broader reportable ranges associated with the real-time PCR assays may translate to fewer repeat tests and lower costs for samples with higher virus loads. The newer assays offer better sensitivity: RTv2 offers a lower detection limit of 20 copies/ml, compared to 40, 48, or 50 copies/ml for ARAUTO (and ARMAN), RTv1, or USCOB (USMON or USCAM), respectively. These lower detection limits may translate to more detectable results. Viral blips are described as transient detections of virus loads that are typically below 200 copies/ml and return to undetectable levels without any change in therapy (15). Increased qualitatively and quantitatively detected results have been noted in laboratories that have switched from the USCAM to the RTv1 assay (11, 24) or RTv2 assay (8). These increases did not translate to increases in mother-to-child transmission (MTCT) rates, nor did they predict virologic failure; however, there is concern that increased occurrences of detectable viremia could result in unnecessary patient anxiety or regimen changes. Further problems could occur with clinical trials where eligibility or treatment failure may be defined based on viral detection. Since differences do exist between the HIV-1 RNA assays, it will be important that guidelines and clinical trial descriptions be updated to reflect these differences. Assay sensitivity across the new real-time PCR assays was evaluated using controls with nominal values of 25 and 50 copies/ml. Better detection rates were noted in citrated controls than in EDTA controls for both the 25-copies/ml (90 to 92% versus 67 to 88%) and 50-copies/ml (96 to 100% versus 71 to 100%) controls for all real-time PCR assays. A significantly reduced detection rate was noted in ARAUTO for the 50-copies/ml EDTA controls, but this was a small data set and would require additional testing to confirm these observations. The specificities of the real-time PCR assays included in this study were very good, with no true false-positive results reported. No evidence of carryover contamination was noted with any of the PCR methodologies, either endpoint or real-time PCR. This was probably a reflection of the routine qual-

ity control measures against cross-contamination in place at the laboratories participating in this study.

Laboratories were asked to collect samples from subjects with a range of virus loads and to limit the number of samples with undetectable virus loads to $<20\%$, in order to permit adequate sample sizes. Some laboratories had to test a large number of HIV-infected samples in order to obtain sufficient numbers of samples with detectable virus loads. Since the lower detection limits of assays differ, the presence or absence of low-level viremia in patient populations can affect concordance rates. Correlation values did vary, ranging from 0.832 to 0.986, but since there was no standardization of the number of samples tested at each virus load across the linear range of the assays, it would not be fair to over-evaluate these data. Low correlations may artifactually result from having an inadequate spread of data across the range of virus loads. The Bland-Altman plots of the same data showed some systematic trends that were more evident in the linearity plots. The target for the 90% limit of agreement was set at $\pm 0.7 \log_{10}$ RNA copies/ml. There were comparisons across assay platforms for which the 90% LOA exceeded this target (ARAUTO versus STD, ARMAN versus US, RTv1 versus STD, and RTv2 versus STD). For purposes of validation, laboratories were instructed to evaluate instances where discordance in correlation testing exceeded $0.5 \log_{10}$ RNA copies/ml. In some cases, the discordance was due to sample mishandling or misreporting of data. Retesting of samples was not always possible, due to sample volume issues, but no major underquantitation was verified as being due to a primer or probe mismatch or an ongoing problem with clinical populations in any of the validation data sets.

In summary, as reported in other studies, differences in \log_{10} recoveries do exist across the assay platforms, but good overall correlations have been noted between endpoint PCR assays and the RTv2 or AR assays (2, 5, 25, 30). Some deficiencies were previously observed with the RTv1 assay but appeared to be mostly resolved with the RTv2 assay (6, 7, 11). However, due to the rapid evolution of HIV-1 in response to selective pressures and the incidence of newly identified recombinant viruses, ongoing surveillance of HIV-1 RNA assay performance will be critical. Questionable results should be retested on a different HIV-1 RNA platform to help resolve potential discordance (4). With respect to low virus titers, donor screening and eligibility testing performed across different platforms may prove challenging, because the lower limits of detection vary across assays. Low titers of HIV-1 RNA may be reported as undetectable by one assay, detectable but less than the cutoff by a second assay, and reported as a quantitative value by a third assay with a lower limit of detection. If detection is used as an eligibility criterion for protocol enrollment, exclusion or inclusion might be biased, depending on the assay used to measure virus load. If laboratories do not routinely report qualitatively detected results, then intermittent low-level viral titers may be misrepresented as virologic failure or misinterpreted as mishandled results. Similarly, if virologic endpoints are placed near or at the limits of detection of different assay platforms, variability in clinical outcomes of trials could be affected by the assay used for virus load testing. Additional studies need to be done to evaluate the effect of using low-copy-number virologic endpoints on clinical outcomes in relation to the assay used for measuring virus load. An understanding of the impact of assay performance is critical, since virus loads are used for clinical management. Assessing performance and determining how certain specifications, such

as accuracy, precision, and lower limits of detection, will impact your results are important for making the right decisions in designing clinical trials and changing treatments.

ACKNOWLEDGMENTS

The Virology Quality Assurance Laboratory acknowledges the work done by all the testing laboratories for helping to generate the data used for these analyses. The testing laboratories included the UNC Retrovirology Laboratory at the University of North Carolina at Chapel Hill (U01 AI068632 and P30 AI50410); the Infectious Disease Laboratory at Children's Memorial Hospital in Chicago, IL; the PHPT Laboratory at FAMS, CMU, in Chiang Mai, Thailand; the CAPRISA Research Laboratory in Durban, South Africa; the KCMC Biotechnology Laboratory in Moshi, Tanzania; The Duke Clinical Trials Unit and Clinical Research Sites (U01 AI069484) and the Center for HIV/AIDS Vaccine Immunology (U01 AI067854), U.S. NIH-funded programs; the CHAVI Core Virology Laboratory at the Duke Human Vaccine Institute in Durham, NC; the Johns Hopkins University HIV-1 Laboratory in Baltimore, MD, supported by NIH funding from U01-AI-068613 and UM1-AI-068613; the Specialized Molecular Diagnostics Unit at the NICD Laboratory in Johannesburg, South Africa; the National Public Health Laboratory in Kingston, Jamaica; the CLS/NHLS Laboratory at the WITS Medical School in Johannesburg, South Africa; the HIV-1-NAT Research Laboratory at the Thai Red Cross Society in Bangkok, Thailand; the Research Institute for Health Sciences Laboratory in Chiang Mai, Thailand; the YRG CARE Infectious Disease Laboratory in Chennai, India; the Centre for Infectious Disease Research in Zambia (CIDRZ) Central Laboratory in Lusaka, Zambia; the Microbiology Research Laboratory at FAM, Khonkaen University in Khonkaen, Thailand; the Botswana—Harvard School of Public Health HIV-1 Research Laboratory (BHHRL) in Gaborone, Botswana; the University of Pittsburgh ACTG-MACS Laboratory in Pittsburgh, PA; the BARC-Lancet Laboratory in Johannesburg, South Africa; the AMPATH Reference Laboratory in Eldoret, Kenya; the UZ-UCSF Central Laboratory in Harare, Zimbabwe; the Grupo Bioquimico/Hospital Italiano Laboratory in Buenos Aires, Argentina; the Walter Reed Project Laboratory in Kericho, Kenya; the National AIDS Research Institute (NARI) Laboratory in Pune, India; the Medical Research Council (MRC)/Uganda Virus Research Council (UVRI) Laboratory in Entebbe, Uganda; the GHESKIO Laboratory in Port-au-Prince, Haiti; the UAB Hospital Molecular Diagnostics Laboratory in Birmingham, AL; the BJ Medical College Laboratory in Pune, India; the Laboratorio de Pasquisa Clinica Evandro Chagas (IPEC) at FIOCRUZ in Rio de Janeiro, Brazil; the JHU Research Project Laboratory at the Queen Elizabeth Central Hospital in Blantyre, Malawi; the Laboratory of Parasite Immunology and Pathology at the University of Puerto Rico School of Medicine in San Juan, PR; and the Joint Clinical Research Center (JCRC) Laboratory in Kampala, Uganda.

We give special thanks to Joe Fitzgibbon for his careful review of the manuscript.

This work was supported by the National Institutes of Health contract HHSN266200500044C, NO1 AI50044 (VQA; principle investigator, J. Bremer).

REFERENCES

- Berger A, et al. 2002. Evaluation of the Cobas AmpliPrep/Cobas AmpliCor HIV-1 Monitor Ultrasensitive test: comparison with the Cobas AmpliCor HIV-1 Monitor test (manual specimen preparation). *J. Clin. Virol.* 25(Suppl. 3):103–107.
- Bourlet T, et al. 2011. HIV-1 load comparison using four commercial real-time assays. *J. Clin. Microbiol.* 49:292–297.
- Brambilla D, et al. 1999. The contribution of assay variation and biological variation to the total variability of plasma HIV-1 RNA measurements. *AIDS* 13:2269–2279.
- Church D, et al. 2011. Comparison of the RealTime HIV-1, COBAS TaqMan 48 v1.0, Easy Q. v1.2, and Versant v3.0 assays for determination of HIV-1 viral loads in a cohort of Canadian patients with diverse HIV-1 subtype infections. *J. Clin. Microbiol.* 49:118–124.
- Crump JA, et al. 2009. Evaluation of the Abbott m2000rt RealTime HIV-1 assay with manual sample preparation compared with the Roche COBAS AmpliPrep/COBAS Amplicor HIV-1 monitor v1.5 using specimens from East Africa. *J. Virol. Methods* 162:218–222.
- Damond F, et al. 2010. Evaluation of an upgraded version of the Roche COBAS AmpliPrep/COBAS TaqMan HIV-1 test for HIV-1 load quantification. *J. Clin. Microbiol.* 48:1413–1416.
- De Bel A, et al. 2010. Corrections of underquantification of human immunodeficiency virus type 1 load with the second version of the Roche COBAS AmpliPrep/COBAS TaqMan assay. *J. Clin. Microbiol.* 48:1337–1342.
- De Boer MGJ, Wessels E, Claas ECJ, Kroon FP. 2010. Potential influence of more-sensitive HIV-1 load detection by the new Roche COBAS AmpliPrep COBAS TaqMan version 2.0 assay on clinical management of HIV-1-positive pregnant women. *J. Clin. Microbiol.* 48:4301–4302.
- Erali M, Hillyard DR. 1999. Evaluation of the Ultrasensitive Roche Amplicor HIV-1 Monitor assay for quantitation of human immunodeficiency virus type 1 RNA. *J. Clin. Microbiol.* 37:792–795.
- Glaubitz J, et al. 2011. Accuracy to 2nd international HIV-1 RNA WHO standard: assessment of three generations of quantitative HIV-1 RNA nucleic acid amplification tests. *J. Clin. Virol.* 50:119–124.
- Gueudin M, et al. 2007. Evaluation of the Roche COBAS TaqMan and Abbott RealTime extraction-quantification systems for HIV-1 subtypes. *J. Acquir. Immune Defic. Syndr.* 44:500–505.
- Holguin A, Aracil B, Alvarez A, Barros C, Soriano V. 2001. Prevalence of human immunodeficiency virus type 1 (HIV-1) non-B subtypes in foreigners living in Madrid, Spain, and comparison of the performances of the Amplicor HIV-1 Monitor version 1.0 and the new automated version 1.5. *J. Clin. Microbiol.* 39:1850–1854.
- Holguin A, Lopez M, Molinero M, Soriano V. 2008. Performance of three commercial viral load assays, Versant human immunodeficiency virus type 1 (HIV-1) RNA bDNA v3.0, COBAS AmpliPrep/COBAS TaqMan HIV-1, and NucliSens HIV-1 EasyQ v1.2, testing HIV-1 non-B subtypes and recombinant variants. *J. Clin. Microbiol.* 46:2918–2923.
- Karasi JC, et al. 2011. High correlation between Roche COBAS AmpliPrep/COBAS TaqMan HIV-1, v2.0 and the Abbott m2000 RealTime HIV-1 assays for quantitation of viral load in HIV-1 B and non-B subtypes. *J. Clin. Virol.* 52:181–186.
- Lee PK, Kieffer TL, Siliciano RF, Nettles RE. 2006. HIV-1 viral load blips are of limited clinical significance. *J. Antimicrob. Chemother.* 57:803–805.
- Mellors JW, et al. 1997. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. *Ann. Intern. Med.* 126:946–954.
- Oliver AR, Pereira SF, Clark DA. 2007. Comparative evaluation of the automated Roche TaqMan RealTime quantitative human immunodeficiency virus type 1 RNA PCR assay and the Roche Amplicor version 1.5 conventional PCR assay. *J. Clin. Microbiol.* 45:3616–3619.
- Pas S, et al. 2010. Performance evaluation of the new Roche COBAS AmpliPrep/COBAS TaqMan HIV-1 test version 2.0 for quantification of human immunodeficiency virus type 1 RNA. *J. Clin. Microbiol.* 48:1195–1200.
- Perrin L, et al. 2006. Multicenter performance evaluation of a new TaqMan PCR assay for monitoring human immunodeficiency virus RNA load. *J. Clin. Microbiol.* 44:4371–4375.
- Rouet F, et al. 2007. Impact of HIV-1 genetic diversity on plasma HIV-1 RNA quantitation. *J. Acquir. Immune Defic. Syndr.* 45:380–388.
- Schumacher W, et al. 2007. Fully automated quantification of human immunodeficiency virus (HIV-1) type 1 RNA in human plasma by the COBAS AmpliPrep/COBAS TaqMan system. *J. Clin. Virol.* 38:304–312.
- Scott LE, et al. 2009. Evaluation of the Abbott m2000 RealTime human immunodeficiency virus type 1 (HIV-1) assay for HIV-1 load monitoring in South Africa compared to the Roche COBAS AmpliPrep/COBAS Amplicor, Roche COBAS AmpliPrep/COBAS TaqMan HIV-1, and BioMerieux NucliSens EasyQ HIV-1 assays. *J. Clin. Microbiol.* 47:2209–2217.
- Sire JM, et al. 2011. Comparative RNA quantification of HIV-1 group M and non-M with Roche COBAS AmpliPrep/COBAS TaqMan HIV-1 v2.0 and Abbott Real-Time HIV-1 PCR assays. *J. Acquir. Immune Defic. Syndr.* 56:239–243.
- Smit E, Bhattacharya S, Osman H, Taylor S. 2009. Increased frequency of HIV-1 viral load blip rate observed after switching from Roche COBAS Amplicor to COBAS TaqMan assay. *J. Acquir. Immune Defic. Syndr.* 51:364–365. (Letter.)
- Ssebgenyi I, et al. 2011. Comparison of the Abbott m2000 HIV-1 Real-

- Time and Roche Amplicor Monitor v1.5 HIV-1 assays on plasma specimens from Rakai, Uganda. *Int. J. STD AIDS* 22:373–375.
26. Swanson P, et al. 2005. Impact of human immunodeficiency virus type 1 (HIV-1) genetic diversity on performance of four commercial viral load assays: LCx HIV-1 RNA quantitative, Amplicor HIV-1 Monitor v1.5, Versant HIV-1 RNA 3.0, and Nuclisens HIV-1 QT. *J. Clin. Microbiol.* 43:3860–3868.
 27. Tang N, et al. 2007. A real-time HIV-1 viral load assay for automated quantitation of HIV-1 RNA in genetically diverse group M subtypes A-H, group O and group N samples. *J. Virol. Methods* 146:236–245.
 28. Triques K, et al. 1999. Efficiencies of four versions of the Amplicor HIV-1 Monitor test for quantification of different subtypes of human immunodeficiency virus type 1. *J. Clin. Microbiol.* 37:110–116.
 29. US Department of Health and Human Services. 2011. Clinical guide for HIV-1/AIDS clinical care. US Department of Health and Human Services, Health Resources and Services Administration, HIV-1/AIDS Bureau, Rockville, MD. <http://hab.hrsa.gov/deliverHIV-1aidsicare/clinicalguide11>.
 30. Van Rensburg EJ, Tait K, Watt A, Schall R. 2011. Comparative evaluation of the Roche COBAS AmpliPrep/COBAS TaqMan HIV-1 version 2 testing using the TaqMan 48 analyzer and the Abbott RealTime HIV-1 assay. *J. Clin. Microbiol.* 49:377–379.
 31. Xu S, et al. 2008. Comparative evaluation of the COBAS AmpliPrep/COBAS TaqMan HIV-1 type 1 test (CAP/CTM) and Versant HIV-1 type 1 RNA 3.0 assay (bDNA) for quantifying HIV-1 type 1 viral loads in China. *AIDS Res. Hum. Retrovir.* 24:1365–1373.
 32. Yen-Lieberman B, et al. 1996. Evaluation of a quality assurance program for quantitation of human immunodeficiency virus type 1 RNA in plasma by the AIDS clinical trials group virology laboratories. *J. Clin. Microbiol.* 34:2695–2701.