JCM
Journals.ASM.org

Monitoring the Quality of HIV-1 Viral Load Testing through a Proficiency Testing Program Using Dried Tube Specimens in Resource-Limited Settings

Shon Nguyen,[a] Artur Ramos,[a] Joy Chang,[a] Bin Li,[b] Vedapuri Shanmugam,[a] Debrah Boeras,[a] John N. Nkengasong,[a] Chunfu Yang,[a] Dennis Ellenberger[a]

Division of Global HIV/AIDS, Center for Global Health,[a] and Division of HIV/AIDS Prevention, National Center for HIV, Hepatitis, STD and TB Prevention,[b] Centers for Disease Control and Prevention, Atlanta, Georgia, USA

HIV-1 viral load (VL) levels are used for monitoring disease progression and antiretroviral therapy outcomes in HIV-infected patients. To assess the performance of laboratories conducting HIV-1 VL testing in resource-limited settings, the U.S. Centers for Disease Control and Prevention implemented a voluntary, free-of-charge, external quality assurance program using dried tube specimens (DTSs). Between 2010 and 2012, DTS proficiency testing (PT) panels consisting of 5 specimens were distributed at ambient temperature to participants. The results from the participants ($n \geq 6$) using the same assay were grouped, analyzed, and graded as acceptable within a group mean ± 3 standard deviations. Mean proficiency scores were calculated by dividing the combined PT scores by the number of testing cycles using a linear regression model. Between 2010 and 2012, the number of participants enrolled increased from 32 in 16 countries to 114 in 44 countries. A total of 78.2% of the participants reported results using 10 different VL assays. The rates of reporting of acceptable results by the participants were 96.6% for the Abbott assay, 96.3% for the Roche Cobas assay, 94.5% for the Roche Amplicor assay, 93.0% for the Biocentric assay, and 89.3% for the NucliSens assay. The overall mean proficiency scores improved over time ($P = 0.024$). DTSs are a good alternative specimen type to plasma specimens for VL PT programs, as they do not require cold chain transportation and can be used on PCR-based assays. Our data suggest that the CDC HIV-1 VL PT program using DTSs positively impacts the testing performance of the participants, which might translate into better and more accurate VL testing services for patients.

High-quality clinical laboratories are urgently needed to sustain global efforts to expand the number of individuals infected with human immunodeficiency virus (HIV) receiving antiretroviral therapy (ART) (1). At the end of 2012, over 35 million people were estimated to be living with HIV, with more than two-thirds of new HIV infections occurring in sub-Saharan Africa (2). Worldwide, ART has been shown to effectively reduce the rates of morbidity and mortality associated with AIDS (3). Quantitation of HIV-1 viral loads (VLs) has become the standard of care for monitoring the response to ART in HIV-infected patients, understanding disease progression, and preventing HIV transmission (4, 5). On the basis of the World Health Organization (WHO) 2010 treatment guidelines, which recommended initiating ART at a CD4 count threshold of 350 cells/mm^3, it was estimated that over 8 million people in low- and middle-income countries were receiving ART in 2011 (6). The number of patients eligible for ART is expected to rise as countries adopt the new 2013 WHO treatment guidelines, which recommend a CD4 count threshold of 500 cells/mm^3 for initiation of ART (1), and the Joint United Nations Programme on HIV/AIDS (UNAIDS) Treatment 2015 plan, which calls for starting 15 million people on ART by 2015 (2). The President's Emergency Plan for AIDS Relief (PEPFAR) blueprint for an AIDS-free generation calls for increased access to VL testing and other technologies to better monitor the effectiveness of ART (7). To meet the demand, many countries, especially in African and Asian regions, with a high HIV prevalence are rapidly expanding their clinical laboratories' capability for HIV VL testing using either commercially available or in-house-developed assays. Access to proficiency testing (PT) panels is an essential component of a comprehensive laboratory external quality assurance (EQA) program that monitors the quality of laboratory testing and accurate reporting of HIV VL testing results (8). In 2010, to ensure the quality of HIV VL testing and services in PEPFAR-supported countries, the Centers for Disease Control and Prevention (CDC) in Atlanta, GA, USA, implemented a voluntary and cost-free PT program for HIV-1 VL quantitation using dried tube specimens (DTSs). Traditional PT panels use liquid plasma specimens, but the cost required to maintain cold chain transportation of these specimens for EQA programs is high (8). Moreover, DTSs are easier to produce than lyophilized materials, which require additional instrumentation. The DTS PT panels offer several advantages over traditional plasma panels: DTSs are noninfectious, are stable at ambient temperature for up to 8 weeks, and do not require special shipping arrangements; hence, the cost for shipment is lower and DTSs can be tested on existing PCR-based HIV-1 VL testing platforms (9). In this report, we present the results of an

Received 25 September 2014 **Returned for modification** 28 October 2014
Accepted 15 January 2015

Accepted manuscript posted online 21 January 2015

Citation Nguyen S, Ramos A, Chang J, Li B, Shanmugam V, Boeras D, Nkengasong JN, Yang C, Ellenberger D. 2015. Monitoring the quality of HIV-1 viral load testing through a proficiency testing program using dried tube specimens in resource-limited settings. J Clin Microbiol 53:1129–1136. doi:10.1128/JCM.02780-14.

Editor: M. J. Loeffelholz

Address correspondence to Chunfu Yang, CYang1@cdc.gov.

doi:10.1128/JCM.02780-14

analysis of the data from an evaluation of the CDC HIV-1 DTS PT program conducted between 2010 and 2012.

MATERIALS AND METHODS

Program enrollment. Between March 2010 and October 2012, all molecular testing laboratories in resource-limited settings, including those supported by PEPFAR and actively performing HIV-1 VL testing, were encouraged to enroll. Laboratories were registered after they electronically submitted a completed enrollment form. Those laboratories which chose to enroll more than one VL assay were welcomed, but each assay was counted as a separate participant. Once enrolled, CDC assigned a unique identification number (ID) to each participant, and only the ID appeared on the results summary report form to protect the participants' confidentiality.

Program materials. Each PT panel package contained two identical sets of five DTSs, one tube of 13 ml reconstitution buffer (NucliSens lysis buffer [bioMérieux, Inc., Hazelwood, MO] or 0.1 M phosphate-buffered saline [PBS], pH 7.4 [Gibco formulation; Invitrogen Corp., Carlsbad, CA]) with its material and safety data sheet (MSDS), instructions, process checklist form, and a result submission form in a Ziploc bag. All packages were shipped at ambient temperature by commercial shipping companies via air cargo and delivered to the participants within an average of 6 days. For each shipment of the PT cycle, participants were instructed to process only one set of DTS PT panels for HIV-1 VL testing and report their results to CDC electronically via email by the established due date. The second set of DTS PT panels in each shipment was provided as a backup, in case the participants had problem during testing of the first set.

DTS PT panels. The DTSs were prepared by diluting inactivated HIV-1 stock strain 97USNG30, a subtype C strain (catalog number 4115; NIH AIDS Reagent Program, Germantown, MD), in 0.1 M PBS buffer containing 2% of green liquid food dye (Kroger, Cincinnati, OH, USA) for visibility purposes. The HIV-1 stock was inactivated by heating at 56°C for 30 min (10). Each PT panel contained RNA at concentrations ranging from 2.0 to 6.0 $\log_{10}$ RNA copies per tube. Twenty microliters of the HIV-1 stock dilutions was dispensed into uncapped 2-ml screw-cap tubes and left at room temperature inside a biosafety cabinet to dry overnight (9, 11). Prior to the panel distribution, the five DTS tubes in each panel were labeled with coded identifiers: (i.e., VL2011-A1 to -A5). Each PT panel in 2012 and 2011 and the panel from the first cycle in 2010 contained one HIV-negative DTS (PBS buffer with 2% food dye and no virus) and four HIV-1-positive DTSs, while the remaining PT panels contained two HIV-negative and three HIV-1 positive DTSs. Overall, 30 HIV-1-positive and 10 HIV-1-negative DTSs were distributed for the PT program between 2010 and 2012. The HIV-1 RNA concentrations varied in the positive DTSs, and codes were randomly changed among panel members during each PT testing cycle.

Panel validation. Multiple sets of DTS PT panels were randomly selected and tested in triplicate in the CDC Molecular Diagnostics Laboratory using four commercial HIV-1 VL assays according to the manufacturers' instructions: the Abbott RealTime HIV-1 test (Abbott Molecular, Des Plaines, IL) (Abbott), the Roche Cobas AmpliPrep/Cobas TaqMan HIV-1 test (Roche Cobas), the Roche Amplicor HIV-1 Monitor (v1.5) test (Roche Diagnostics, Indianapolis, IN) (Roche Amplicor), and the NucliSens EasyQ HIV-1 test (bioMérieux, France) (NucliSens). Newly prepared DTS PT panels were validated by verifying that the intrarun HIV-1 VL results for the triplicate samples of each panel member fell within a range of 0 to 0.5 $\log_{10}$ copy/ml.

HIV-1 VL assay. The DTSs were reconstituted by adding 1.1 ml of reconstitution buffer to each tube, followed by 10 s of mixing on a vortex instrument. The NucliSens lysis buffer was used as the reconstitution buffer for the first three PT testing cycles in 2010, and 0.1 M PBS buffer was used for the subsequent five PT testing cycles in 2011 and 2012. The reconstituted DTSs were tested immediately by following the recommendations for HIV VL testing of the manufacturers of the commercial assays for clinical specimens or the participant-developed laboratory standard operating procedures for in-house assays.

Data analysis. The reported data were evaluated with the respect to the ability of the participants to obtain VL testing results that reproduce the known concentrations of HIV-1 RNA present in each DTS PT panel member. The results reported from the participants were grouped and analyzed by type of VL assay to avoid possible variability of the results among different types of VL assays (12–14). These groups were formed by a minimum of six participants using the same assay in each PT testing cycle. For each group, a box-and-whisker plot (15) was used as pretreatment of the data to identify outliers for each DTS PT member. The outliers were defined as data points that fell below the first quartile (25th percentile) and above the third quartile (75th percentile) from the median. These outlier results were excluded from the calculation of the group mean, standard deviation (SD), and group coefficient variation (CV). The limits of acceptability for each panel member were assigned in a range within the group mean $\pm$ 3 SDs. All data points for DTS PT results that fell within or outside this range were graded as acceptable or unacceptable, respectively. Participants received a passing score if a minimum of four of the five panel members (80.0%) yielded an acceptable grade in a given PT testing cycle, on the basis of the criteria of the WHO Regional Office for Africa accreditation process (16). The results from participants who reported fewer than six PT results using a single assay or who used in-house VL assays in a given PT testing cycle were not graded, and results from these assays were combined into a group labeled "other assays." The graded and ungraded results for each panel member and its group mean, SD, and CV were recorded on the results summary report form that was sent electronically to the appropriate participants within 4 weeks. Each participant was instructed to review the results summary report form for comparison of its results with those of its peers, investigate potential sources of error, and initiate corrective action and a request for technical assistance from CDC, if needed. To monitor the participants' performance over time, each participant's overall mean proficiency score was calculated by adding all the individual PT panel scores and dividing by the total number of cycles in which the participants participated and analyzed using a linear regression model.

RESULTS

Program participation. In 2010, when the PT program was initiated, 32 participants in 16 countries participated in the first cycle. However, by the end of 2012, the number of participants had increased to 114 in 44 countries, and more than 95% of participants continued in the program after initial enrollment (Fig. 1). Approximately 6% of participants in 2010 and 2011 and 10% of participants in 2012 had enrolled two VL assays into the DTS VL PT program (data not shown).

Program response. Of the 665 packages containing 6,650 DTS tubes that were shipped, results from 520 packages (78.2%; range, 71.9% to 85.2%) for eight PT testing cycles were reported to CDC. These included 23 of the 32 packages for the VL2010-A PT cycle, 46 of the 54 for the VL2010-B PT cycle, 48 of the 58 for the VL2010-C PT cycle, 61 of the 82 for the VL2011-A PT cycle, 78 of the 107 for the VL2011-B PT cycle, 91 of the 110 for the VL2011-C PT cycle, 88 of the 108 for the VL2012-A, and 85 of the 114 for the VL2012-B PT cycle (Fig. 1). Over the entire eight testing cycles, three participants reported using the second set of PT panels after experiencing an electrical power failure during their VL testing of the first set of PT panels. Participants who did not report results provided the following reasons: a lack of reagent kits, expired reagents, instrument issues, a lack of persons to perform the tests, insufficient electric power, interruption because of laboratory facility renovation, shipping issues, issues with clearance through their country's customs, and improper handling of PT panel pack-

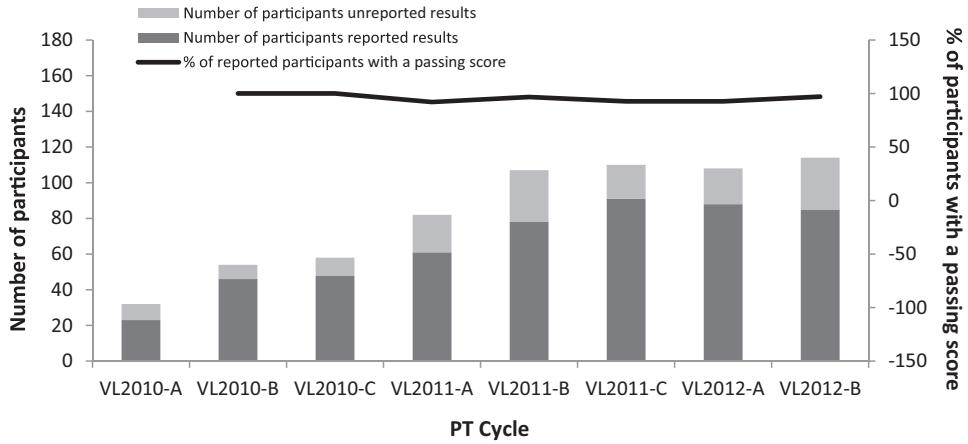

FIG 1 Summary of the number of participants who obtained a passing score during eight PT cycles between 2010 and 2012. The number of participants reporting results and the percentage of those participants with a passing score (80%) at each PT testing cycle are presented.

ages. Ten different VL assays were used to obtain the reported results: the Abbott RealTime HIV-1 test (Abbott), the Roche Cobas AmpliPrep/Cobas TaqMan HIV-1 test (Roche Cobas), the Roche Amplicor HIV-1 Monitor (v1.5) test (Roche Amplicor), the NucliSens EasyQ HIV-1 test (NucliSens), the Biocentric generic HIV viral load test (Biocentric, Bandol, France) (Biocentric), the Exavir Load-Cavidi (v3) test (Cavidi AB, Uppsala, Sweden), and in-house VL assays. The in-house VL assays were the Gag-Sybr green assay, the Cobas TaqMan with HighPure system, the HighPure system viral nucleic acid kit, and the Artus HIV Virus-1 RG reverse transcription-PCR kit. Among the 10 VL assays used, the Roche Cobas and Roche Amplicor tests were those used the most often in this PT program, being used by a range of from 15% to 41% of participants, followed by the Abbott test, which was used by a range of from 14% to 24% of participants (Fig. 2).

Data analysis. Due to undetectable viral RNA from the VL2011-A3 panel by the Abbott and Roche Amplicor assays, only 29 matched HIV-1-positive DTS members were analyzed for the comparison of the PT results (Table 1; see Fig. 4). Among the different HIV-1 VL assays, the Abbott, Roche Cobas, and Roche Amplicor assays indicated good reproducibility and agreement, as a regression coefficient (R^2) value of 0.97 was observed between the Abbott and Roche Cobas assays, and an R^2 value of 0.95 was observed between the Abbott and Roche Amplicor and between the Roche Cobas and Roche Amplicor assays; one sample, however, was reported to have a VL below the limit of detection by the Roche Amplicor assay (Fig. 3). The correlation between participants using the Biocentric, NucliSens, and other assays was not determined because the data obtained using these assays were not available for the 29 positive DTS members of the PT panels for the time period of this study (Table 1).

The overall interparticipant variability of HIV-1 RNA quantification by the Abbott and Roche Cobas tests was consistently low, with the group SD values being below 0.5 and the CV values being below 16.0% for each platform. The group SD values observed for the Roche Amplicor assay were below 0.8 for all panels except

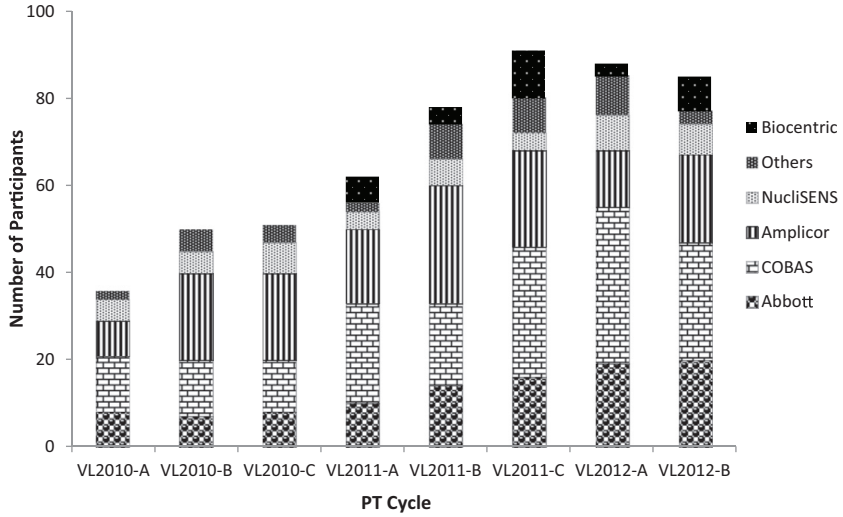

FIG 2 Number of different viral load assays used by participants over time (eight PT cycles). Participants reported using the Abbott RealTime HIV-1 test, the Cobas AmpliPrep/Cobas TaqMan HIV-1 test, the Amplicor HIV-1 Monitor (v1.5) test, the NucliSens EasyQ HIV-1 test, the Biocentric test, and other assays for the eight PT cycles between 2010 and 2012.

TABLE 1 Number of participants in individual PT panels with acceptable and unacceptable grades for HIV-1 RNA VL testing by the Abbott, Roche Cobas, Roche Amplicor, Biocentric, and NucliSens assays[a]

Yr	Sample ID	Abbott (1.6)[b]					Roche Cobas (1.6)					Roche Amplicor (2.6)					Biocentric (2.5)					NucliSens (1.4)				
			LA		*n*			LA		*n*			LA		*n*			LA		*n*			LA		*n*	
		n	L	H	A	U	*n*	L	H	A	U	*n*	L	H	A	U	*n*	L	H	A	U	*n*	L	H	A	U
2010	B1	7	2.1	4.6	7	0	13	2.8	4.2	13	0	20	1.7	5.1	19	1										
	B2	7	5	6.7	7	0	13	4.9	7.1	13	0	20	4.8	6.6	19	1										
	B3	7	NA	NA	7	0	13	NA	NA	13	0	20	NA	NA	20	0										
	B4	7	3.7	5.6	7	0	13	3.9	5.6	12	1	20	3	6.1	20	0										
	B5	7	NA	NA	7	0	13	NA	NA	13	0	20	NA	NA	19	1										
	C1	8	3.4	5	8	0	12	3.4	5.3	12	0	20	2.7	5.5	19	1	7	3.9	6.3	7	0					
	C2	8	3	5	8	0	12	3.5	5.1	12	0	20	3.1	5.3	16	4	7	4	6.1	7	0					
	C3	8	NA	NA	8	0	12	NA	NA	12	0	20	NA	NA	20	0	7	NA	NA	7	0					
	C4	8	4.3	6.2	8	0	12	4.9	6.2	12	0	20	3.8	6.6	18	2	7	5.2	7.1	7	0					
	C5	8	NA	NA	8	0	12	NA	NA	12	0	20	NA	NA	20	0	7	NA	NA	7	0					
2011	A1	10	NA	NA	10	0	23	NA	NA	23	0	17	NA	NA	15	2										
	A2	10	3.2	4.1	8	2	23	3	4.9	22	1	17	2.9	5.1	15	2										
	A3	10	NA	NA	10	0	23	1.5	2.4	21	2	17	NA	NA	17	0										
	A4	10	3.2	5.6	10	0	23	3.5	6	23	0	17	3.2	6.5	17	0										
	A5	10	1.2	2.7	10	0	23	1.5	2.4	21	2	17	NA	NA	17	0										
	B1	14	1.4	3.9	14	0	19	2.3	3.5	19	0	27	1.6	3.6	26	1	8	1.6	4.9	8	0					
	B2	14	4.3	5.7	14	0	19	4.5	5.8	18	1	27	3.7	6.6	25	2	8	3.7	7.3	8	0					
	B3	14	4.3	5.7	14	0	19	4.5	5.8	19	0	27	3.4	6.8	26	1	8	3.7	7.3	8	0					
	B4	14	NA	NA	14	0	19	NA	NA	19	0	27	NA	NA	27	0	8	NA	NA	8	0					
	B5	14	1.7	3.7	14	0	19	2.3	3.4	19	0	27	1.9	3.5	26	1	8	1.1	5.5	7	1					
	C1	16	1.7	3.1	15	1	30	2.4	2.9	25	5	22	1.3	3.8	21	1	8	2.1	3.7	7	1					
	C2	16	NA	NA	15	1	30	NA	NA	28	2	22	NA	NA	22	0	8	NA	NA	8	0					
	C3	16	4	6.1	16	0	30	4.5	5.9	27	3	22	2.5	7.1	21	1	8	4.9	5.8	6	2					
	C4	16	3.3	4.6	15	1	30	3.5	4.7	28	2	22	2.4	5.4	21	1	8	3.2	5.6	8	0					
	C5	16	4.7	5.6	12	4	30	3.9	6.1	28	2	22	3.5	6.7	20	2	8	4.6	6.7	8	0					
2012	A1	19	3.7	5.7	19	0	36	4.6	5.9	34	2	13	3.9	6.4	12	1	9	4.8	6	6	3	8	3.8	6.2	8	0
	A2	19	1.7	3.4	18	1	36	2.2	3.7	34	2	13	1.1	4.9	13	0	9	2	4.3	8	1	8	2.3	3.6	8	0
	A3	19	3.7	5.7	19	0	36	4.6	5.9	34	2	13	4.4	6	10	3	9	4.5	6.5	7	2	8	4	6	8	0
	A4	19	3	4.6	19	0	36	3.6	5	35	1	13	3.5	4.7	10	3	9	3.6	5.5	7	2	8	3.2	4.9	7	1
	A5	19	NA	NA	19	0	36	NA	NA	36	0	13	NA	NA	12	1	9	NA	NA	8	1	8	NA	NA	7	1
	B1	20	2.4	4.5	19	1	27	3.2	4.9	25	2	20	2.1	5.5	19	1	8	2.6	5.6	8	0	7	3	4.8	6	1
	B2	20	NA	NA	20	0	27	NA	NA	27	0	20	NA	NA	18	2	8	0	0	8	0	7	0	0	6	1
	B3	20	1.2	2.7	15	5	27	1.2	3.1	27	0	20	0	4.7	20	0	8	0	6.5	8	0	7	1.5	3.4	6	1
	B4	20	2.2	4.7	20	0	27	3.2	4.7	27	0	20	2.3	5.1	18	2	8	2.9	5.5	8	0	7	3.2	4.7	6	1
	B5	20	3.3	6.1	20	0	27	4.4	5.8	27	0	20	2.4	7.2	19	1	8	3.8	7.2	7	1	7	5.2	5.4	5	2
Total no. (%)		470			454 (96.6)	16 (3.4)	800			770 (96.3)	30 (3.8)	695			657 (94.5)	38 (5.5)	200			186 (93)	14 (7)	75			67 (89.3)	8 (10.7)

[a] *n*, number of participants; LA, limits of acceptability in $\log_{10}$ number of copies/ml; L, low limit of acceptability (mean − 3 SDs) in $\log_{10}$ number of copies/ml; H, high limit of acceptability (mean + 3 SDs) in $\log_{10}$ number of copies/ml; A, number of samples with an acceptable result; U, number of samples with an unacceptable result; NA, not applicable (results that were negative or below the limit of detection).

[b] Values in parentheses are the assay limit of detection in $\log_{10}$ number of copies/ml.

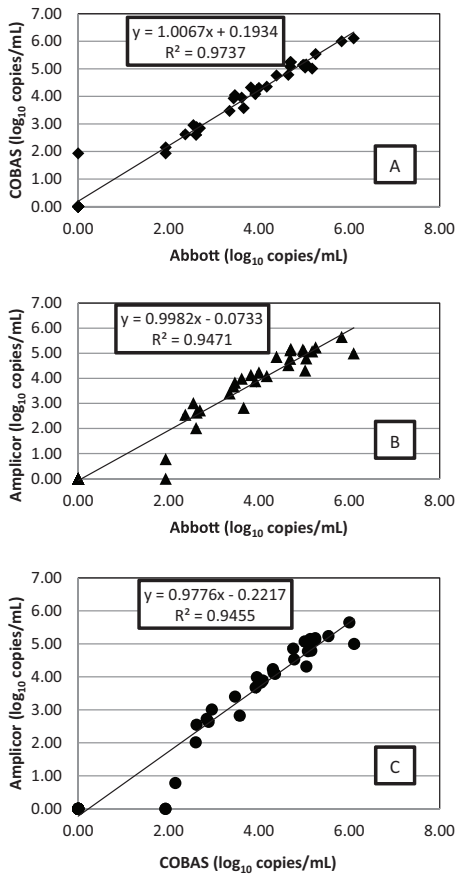

FIG 3 Correlations between viral load results measured by different assays. The group means ($\log_{10}$ number of copies/ml) were plotted between two different assays: Abbott versus Roche Cobas (A), Abbott versus Roche Amplicor (B), and Roche Cobas versus Roche Amplicor (C). The trend line was determined with Microsoft Excel software (linear regression model) and was used to calculate the linear regression formula and R^2 values.

VL2010-A1 and VL2012-B3, for which group SD values were 0.98 and 1.29, respectively. The group CV values for the Roche Amplicor assay were below 21.0% for all panels except VL2012-B3, which had a CV value of 164.9% (Fig. 4). Despite the fact that only 6 to 9 participants used the Biocentric or NucliSens assay, the majority of SD and CV values were below 0.7 and 16.7%, respectively (data not shown). Because less than six participants used other assays, the CV values varied over a wide range, from a low of 2.3% to a high of 29.2%, and the SD values varied from 0.09 to 1.28 (data not shown).

Results grading. Among the 520 reported sets of results, 448 (86.2%) were graded (Table 1). Overall, more than 90% of participants received passing scores in every PT testing cycle analyzed (Fig. 1). Moreover, greater than 94% of results from participants that used the Abbott, Roche Cobas, and Roche Amplicor assays fell within the limit of acceptability (group mean ± 3 SDs) for all panel members. The first PT cycle, VL2010-A, was a pilot event and was not graded. The highest overall rates of acceptable result were from participants using the Abbott (96.6%) and Roche Cobas (96.3%) assays, followed by those using the Roche Amplicor assay (94.5%). Because less than 10 participants used the Biocentric and NucliSens assays, interparticipant comparison of acceptable results was reduced, and the rates of acceptable results were 93.0% for Biocentric and 89.3% for NucliSens (Table 1).

The majority of outliers were graded as unacceptable results, regardless of the type of assay, except for one outlier for a DTS member (VL2010-B4) using the Abbott assay and one outlier for a DTS member (VL2012-B3) using the Roche Cobas assay. These two outliers were graded acceptable since they fell within the acceptable ranges. Among the results of the users of the Abbott assay, 3.4% (16/470) of the results from 12 participants were unacceptable. Among them, 10 participants had one unacceptable result in one PT cycle which resulted in passing that PT cycle, while two participants had two unacceptable results in one PT cycle which resulted in failing that PT cycle. Among the Roche Cobas assay users, 3.8% (30/800) of the results from 15 participants were unacceptable. Among these 15 participants, 9 had one unacceptable

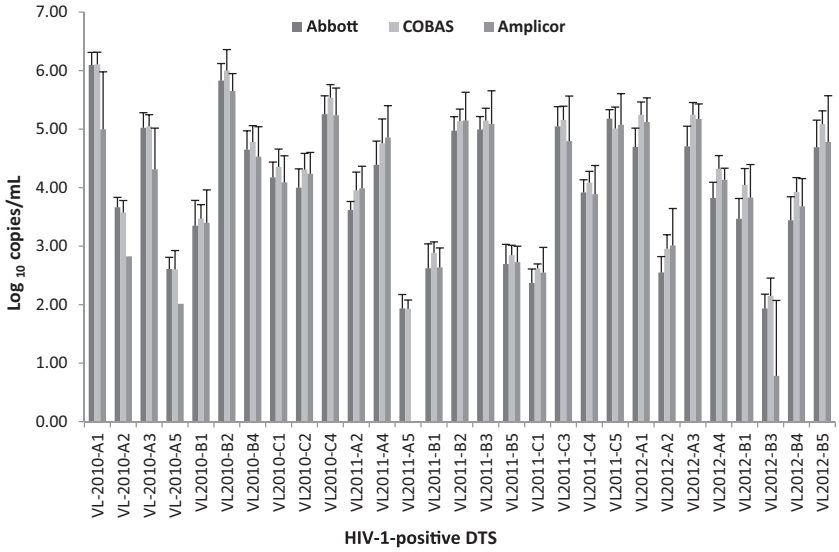

FIG 4 Interassay variability over time. The group mean VL ($\log_{10}$ number of copies/ml) and SD (illustrated by whiskers) for each HIV-1-positive sample determined by the Abbott, Roche Cobas, and Roche Amplicor assays for the eight PT cycles are shown.

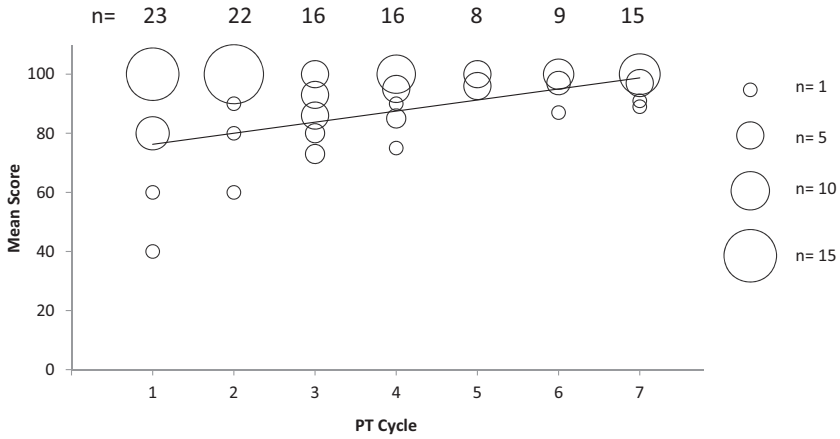

FIG 5 Performance of participants according to the number of PT cycles in which they participated. The circle size corresponds to the number of participants. A linear regression for trend was used to analyze the performance of the participants according to the number of PT cycles in which they participated. The seven PT cycles had 23, 22, 16, 16, 8, 9, and 15 participants, respectively, that reported viral load results.

result and 6 had two or more unacceptable results, resulting in those participants failing those PT cycles. Despite the wider acceptability ranges calculated for the Roche Amplicor assay than the Abbott and Roche Cobas assays, 5.5% (38/695) of the results from 15 participants were unacceptable. Among these 15 participants, 9 had one unacceptable result, while 6 participants had two or more unacceptable results, leading those participants to fail the PT score evaluations.

The reasons for unacceptable results (4.3%, 84/1,965) reported from participants using the Abbott, Roche Cobas, and Roche Amplicor assays included false-negative results ($n = 26$), false-positive results ($n = 7$), results outside the limits of acceptability ($n = 47$), and invalid results ($n = 4$) that indicated possible sample mix-up during sample processing. These were observed randomly across these three different assays (data not shown).

To study the impact of the CDC VL PT on VL testing performance over time, we analyzed the trend for the proficiency scores for all graded participants using a linear regression model. The results indicated improved performance over time ($P = 0.024$): overall mean proficiency scores increased as the number of cycles in which participants participated increased, regardless of the VL assays that they used. The overall mean proficiency score increased from 92.0% for participants who participated in one PT testing cycle to 98.0% for participants who participated in all seven PT cycles evaluated (Fig. 5).

DISCUSSION

From the time of introduction of the CDC HIV-1 VL PT program for HIV-1 VL quantitation in 2010, the number of participants tripled by 2012, reflecting the need for this program and the interest of clinical laboratories in monitoring and improving the quality of HIV-1 VL testing for patients. Furthermore, this PT program highlights the importance of monitoring the quality of laboratory tests, as inaccurate VL results might have been reported to clinicians, a possibility inferred from the several unacceptable PT results found in this study. The data analyzed in this study demonstrated the good reproducibility of the results of different PCR-based VL assays among multiple participants. To grade the results of the second PT testing cycle, VL2010-B, the results obtained by a group of six or more participants using the same assay was initially chosen on the basis of the SD and CV values for a group of six participants and a CDC laboratory using the Abbott assay in the first analysis. The SD and CV values for that group of 6 participants were similar to those for a group of 13 participants using the Roche Cobas assay and lower than those for a group of 20 participants using the Roche Amplicor assay (Table 1). The Abbott assay results from the CDC laboratory were added in the analysis of the data from the Abbott assay to increase the sample size from six to seven. To maintain consistency, a group of six participants using the same assay was continuously used for all assays for the onward data analysis.

The overall proficiency scores of participants indicated improved performance over time for all PCR-based VL assays (Fig. 5). Among the participants with overall scores lower than 90%, the majority of the participants used the Roche Amplicor assay, followed by the Roche Cobas and Abbott assays. This finding may be a reflection of the overall interparticipant and interassay variability among these three widely used assays. There are limitations to this VL PT program. First, the DTSs cannot be tested by enzyme activity-based assays, such as the Cavidi assay, as HIV reverse transcriptase activity is inactivated during DTS preparation, which includes a heat inactivation and drying process, and shipment at ambient temperature (9, 10). A previous study using dried plasma specimens containing HIV-1 also revealed the noninfectious nature of the dried plasma specimens (17). Thus, the participants using the Cavidi assay were notified and excluded from the program. Second, the single subtype C viral strain used for the PT program may not fully reflect the subtype-specific VL quantitation variations that VL assays may have.

With the surge in PEPFAR funding, along with the increased demand for HIV VL testing, several sub-Saharan African countries have adopted automated VL assays, such as the Roche Cobas and Abbott assays, to improve testing quality and turnaround times. Therefore, the number of participants using the Roche Cobas and Abbott assays has doubled since the program was initiated. The data show that the interassay variability for these assays was comparable to the intra-assay variability. The CV values for these assays were between 3.0% and 15.9%, which was less than those for the Roche Amplicor assay, which had CV values between 3.0% and 21.0%. The higher CV for the results from the Roche

Amplicor assay group than for the results for the automated assays could be explained by the manual handling of multiple steps of the testing procedure. Other studies have also observed that fully automated high-throughput assays, such as the Abbott and the Roche Cobas assays, show smaller variability than manual assays, such as the Roche Amplicor assay (18). Despite the small number of data points provided by participants using the Biocentric and NucliSens assays, the intra-assay variability was below 17%, which also demonstrated good reproducibility by these semiautomated assays (data not shown).

We noticed that only one participant using the Roche Amplicor assay had repeated failed PT scores on two PT testing cycles from 2010 to 2012. The Roche Amplicor assay was discontinued and was withdrawn from the market at the end of 2012. Nevertheless, this low rate of repeated failures may reflect the possibility that the participants have effectively used the CDC PT program to monitor and improve their laboratory VL testing performance. The reasons for unacceptable grades were scattered across different participants and PT testing cycles for all assay methods. None of the participants reported problems with the sample reconstitution buffer after NucliSens lysis buffer was replaced by 0.1 M PBS buffer as the constitutional buffer for DTSs from the fourth to the eighth PT testing cycles in 2011 and 2012. Also, no difference in VL results was noted when either reconstitution buffer was used (data not shown). These results further validate our decision to switch from NucliSens lysis buffer to the commonly used 0.1 M PBS to reduce the cost of PT panel production, as our previous evaluation study indicated (9).

The CDC PT program has been monitoring individual participant performance trends on the basis of the overall mean proficiency scores as well as trends in issues contributing to unacceptable results and has been providing technical assistance to individual participants to address specific issues. The PT program is also providing a performance-monitoring tool to laboratory managers, allowing implementation of appropriate corrective actions when necessary according to the participant's performance. Considering the many challenges that participants are facing in resource-limited countries and the need to increase reporting of PT results, the participants were allowed to use the second set of PT panels for troubleshooting and investigating issues indicated in their results summary report. A website for this PT program is currently under development. At this website, the participants will be able to access answers to frequently asked questions, check the results due date, submit their results online, and receive the results summary report for their participation to minimize late reporting of results and improve data collection, analysis, and reporting.

In summary, the CDC HIV-1 VL PT program using DTSs has provided an avenue for VL testing laboratories in PEPFAR-supported countries to periodically monitor and evaluate the quality of their VL testing services. The overall PT program analysis reported here indicates that the longer that the participants participate in the PT program, the higher that their overall performance scores are. This improved performance might translate into better and more accurate VL testing services for patients, which might further result in an increase in the efficacy of care and treatment programs and an increase in the effectiveness of HIV prevention programs.

ACKNOWLEDGMENTS

We thank Guoqing Zhang and John Nahabedian from the Centers for Disease Control and Prevention for their assistance with preparation of the figures and to all participants who contributed their data included in this report.

This research has been supported by the President's Emergency Plan for AIDS Relief (PEPFAR) through the Centers for Disease Control and Prevention.

The findings and conclusions in this paper are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention. The use of trade names is for identification only and does not constitute endorsement by the U.S. Department of Health and Human Services, the Public Health Service, or the Centers for Disease Control and Prevention.

REFERENCES

1. **WHO.** 2013. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. WHO, Geneva, Switzerland.
2. **Joint United Nations Programme on HIV/AIDS.** 2013. Global report: UNAIDS report on the global AIDS epidemic 2013. Joint United Nations Programme on HIV/AIDS, Geneva, Switzerland.
3. **WHO.** 2014. Global update on the health sector response to HIV. WHO, Geneva, Switzerland.
4. **Centers for Disease Control and Prevention.** 2001. Guidelines for laboratory test result reporting of human immunodeficiency virus type 1 ribonucleic acid determination. Recommendations from a CDC working group. MMWR Recomm Rep **50**(RR-20)**:**1–12. http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5020a1.htm.
5. **Gunthard HF, Aberg JA, Eron JJ, Hoy JF, Telenti A, Benson CA, Burger DM, Cahn P, Gallant JE, Glesby MJ, Reiss P, Saag MS, Thomas DL, Jacobsen DM, Volberding PA.** 2014. Antiretroviral treatment of adult HIV infection: 2014 recommendations of the International Antiviral Society—USA Panel. JAMA **312:**410–425. http://dx.doi.org/10.1001/jama.2014.8722.
6. **WHO.** 2010. Antiretroviral therapy for HIV infection in adults and adolescents: recommendations for a public health approach, 2010 revision. WHO, Geneva, Switzerland.
7. **The President's Emergency Plan for AIDS Relief.** 2012. PEPFAR blue print: creating an AIDS-free generation. Office of the Global AIDS Coordinator, The President's Emergency Plan for AIDS Relief, Washington, DC.
8. **Lin HJ, Pedneault L, Hollinger FB.** 1998. Intra-assay performance characteristics of five assays for quantification of human immunodeficiency virus type 1 RNA in plasma. J Clin Microbiol **36:**835–839.
9. **Ramos A, Nguyen S, Garcia A, Subbarao S, Nkengasong JN, Ellenberger D.** 2013. Generation of dried tube specimen for HIV-1 viral load proficiency test panels: a cost-effective alternative for external quality assessment programs. J Virol Methods **188:**1–5. http://dx.doi.org/10.1016/j.jviromet.2012.11.036.
10. **Tjotta E, Hungnes O, Grinde B.** 1991. Survival of HIV-1 activity after disinfection, temperature and pH changes, or drying. J Med Virol **35:**223–227. http://dx.doi.org/10.1002/jmv.1890350402.
11. **Parekh BS, Anyanwu J, Patel H, Downer M, Kalou M, Gichimu C, Keipkerich BS, Clement N, Omondi M, Mayer O, Ou CY, Nkengasong JN.** 2010. Dried tube specimens: a simple and cost-effective method for preparation of HIV proficiency testing panels and quality control materials for use in resource-limited settings. J Virol Methods **163:**295–300. http://dx.doi.org/10.1016/j.jviromet.2009.10.013.
12. **Butler IF, Pandrea I, Marx PA, Apetrei C.** 2007. HIV genetic diversity: biological and public health consequences. Curr HIV Res **5:**23–45. http://dx.doi.org/10.2174/157016207779316297.
13. **Lin HJ, Myers LE, Yen-Lieberman B, Hollinger FB, Henrard D, Hooper CJ, Kokka R, Kwok S, Rasheed S, Vahey M.** 1994. Multicenter evaluation of quantification methods for plasma human immunodeficiency virus type 1 RNA. J Infect Dis **170:**553–562. http://dx.doi.org/10.1093/infdis/170.3.553.
14. **Yen-Lieberman B, Brambilla D, Jackson B, Bremer J, Coombs R, Cronin M, Herman S, Katzenstein D, Leung S, Lin HJ, Palumbo P, Rasheed S, Todd J, Vahey M, Reichelderfer P.** 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.

15. **Tukey JW.** 1977. Exploratory data analysis, 1st ed. Pearson, London, United Kingdom.
16. **Gershy-Damet GM, Rotz P, Cross D, Belabbes el H, Cham F, Ndihokubwayo JB, Fine G, Zeh C, Njukeng PA, Mboup S, Sesse DE, Messele T, Birx DL, Nkengasong JN.** 2010. The World Health Organization African region laboratory accreditation process: improving the quality of laboratory systems in the African region. Am J Clin Pathol **134:**393–400. http://dx.doi.org/10.1309/AJCPTUUC2V1WJQBM.
17. **Lloyd RM, Jr, Burns DA, Huong JT, Mathis RL, Winters MA, Tanner M, De La Rosa A, Yen-Lieberman B, Armstrong W, Taege A, McClernon DR, Wetshtein JL, Friedrich BM, Ferguson MR, O'Brien W, Feorino PM, Holodniy M.** 2009. Dried-plasma transport using a novel matrix and collection system for human immunodeficiency virus and hepatitis C virus virologic testing. J Clin Microbiol **47:**1491–1496. http://dx.doi.org/10.1128/JCM.02354-08.
18. **Senechal B, James VL.** 2012. Ten years of external quality assessment of human immunodeficiency virus type 1 RNA quantification. J Clin Microbiol **50:**3614–3619. http://dx.doi.org/10.1128/JCM.01221-12.