Comparative Performances of HIV-1 RNA Load Assays at Low Viral Load Levels: Results of an International Collaboration

Swenson, L C; Cobb, B; Geretti, A M; et al; Böni, J

Abstract: Low-level viremia during antiretroviral therapy and its accurate measurement are increasingly relevant. Here, we present an international collaboration of 4,221 paired blood plasma viral load (pVL) results from four commercial assays, emphasizing the data with low pVL. The assays compared were the Abbott RealTime assay, the Roche Amplicor assay, and the Roche TaqMan version 1 and version 2 assays. The correlation between the assays was 0.90 to 0.97. However, at a low pVL, the correlation fell to 0.45 to 0.85. The observed interassay concordance was higher when detectability was defined as 200 copies/ml than when it was defined as 50 copies/ml. A pVL of 100 to 125 copies/ml by the TaqMan version 1 and version 2 assays corresponded best to a 50-copies/ml threshold with the Amplicor assay. Correlation and concordance between the viral load assays were lower at a low pVL. Clear guidelines are needed on the clinical significance of low-level viremia.

DOI: https://doi.org/10.1128/JCM.02461-13

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-100949

Originally published at:
Comparative Performances of HIV-1 RNA Load Assays at Low Viral Load Levels: Results of an International Collaboration


Low-level viremia during antiretroviral therapy and its accurate measurement are increasingly relevant. Here, we present an international collaboration of 4,221 paired blood plasma viral load (pVL) results from four commercial assays, emphasizing the data with low pVL. The assays compared were the Abbott RealTime assay, the Roche Amplicor assay, and the Roche TaqMan version 1 and version 2 assays. The correlation between the assays was 0.90 to 0.97. However, at a low pVL, the correlation fell to 0.45 to 0.85. The observed interassay concordance was higher when detectability was defined as 200 copies/ml than when it was defined as 50 copies/ml. A pVL of ~100 to 125 copies/ml by the TaqMan version 1 and version 2 assays corresponded best to a 50-copies/ml threshold with the Amplicor assay. Correlation and concordance between the viral load assays were lower at a low pVL. Clear guidelines are needed on the clinical significance of low-level viremia.

The level of HIV-1 RNA in blood plasma (viral load) is arguably the most important surrogate marker in the treatment of HIV infection (1). For over a decade, the endpoint PCR-based Roche Cobas Amplicor HIV-1 Monitor test (versions 1 and 1.5) (Amplicor) (2) was a widely used viral load assay and was in use in both clinical trials and routine practice. Many clinical trials have used a viral load considered to be “undetectable” by the Amplicor assay (i.e., one that is <50 HIV-1 RNA copies/ml) as the endpoint for measuring the efficacy of antiretroviral therapy. Various guidelines have adopted the same threshold as the goal of therapy for all patients, regardless of previous treatment experience (1, 3–5). After many years of using the Amplicor test, the real-time PCR-based Roche Cobas AmpliPrep/Cobas TaqMan real-time quantitative human immunodeficiency virus type 1 (HIV-1) test (TaqMan) began to replace the Amplicor test in early 2008. There are two versions of the TaqMan assay; version 2 has replaced version 1 across Europe, Asia, and North America. The TaqMan versions 1 and 2 assays have lower limits of quantitation of 40 and 20 copies/ml, respectively (6). Other assays based on real-time PCR technology have also been implemented in clinical settings, including the Abbott RealTime HIV-1 assay (RealTime) (Abbott Molecular, Inc.) and the Qiagen artus HIV-1 QS-RQ assay (7, 8), which have lower limits of quantitation of 40 and 45 copies/ml, respectively (9).

Although generally these assays tend to give roughly comparable results (10), their manufacturers have reported that variation and error tend to increase at the lower limits of quantitation of the assays. Furthermore, there have been reports that real-time PCR-based assays tend to identify more patients as having low-level viremia than does the Amplicor test (10–15). Multiple studies have documented variation between these assays when measuring samples with low viral load levels (6, 13, 15–21).

Therapy outcomes have been shown to be influenced by low-level viremia in some studies (16, 22–28) but not others (29–32), in part reflecting differences in the measured outcomes. However, the overall clinical relevance of low-level viremia remains controversial. Accordingly, some clinical guidelines (4, 33) have set higher thresholds (~200 copies/ml) for defining virologic failure, while others (1) are more conservative and recommend that patients with lower levels of viremia be reviewed due to a potential risk of virological rebound.

A significant clinical question has been to determine the relevance of low-level viremia. In order to properly address this question, an assessment of the fundamental agreement of the various assays for defining low-level viremia must be undertaken. Given that the viral load assays mentioned above differ in their reported lower limits of quantitation and detection (LLQ and LLD, respectively) (34), and given the lack of definitive guidelines on the management of low-level viremia, our aim was to assess the basic comparability of various viral load assays at these lower viral loads. Thus, we established a large international collaborative group to assemble comparisons of the Amplicor 1.5, RealTime, TaqMan version 1, and TaqMan version 2 viral load assays at low viral loads.
MATERIALS AND METHODS

A total of 14 sites from Europe, North America, and Africa contributed 4,221 paired viral load results obtained by different assays performed using the same sample. The numbers of data points in the comparisons were 1,384 (Amplicor versus TaqMan version 1), 365 (Amplicor versus TaqMan version 2), 827 (RealTime versus TaqMan version 1), 1,230 (RealTime versus TaqMan version 2), and 415 (TaqMan version 1 versus TaqMan version 2). The subtype data, when provided, were collected independently at each site.

FIG 1

Comparisons among viral load assays. (A and B) Shown is the correlation of the TaqMan assays with either the Amplicor assay (A) or the RealTime assay (B). Comparisons involving TaqMan version 1 are shown with filled gray circles, while those involving TaqMan version 2 are shown with open squares. The $R^2$ values for all correlations are shown in the legends. (C to F) The bottom four panels show Bland-Altman difference plots comparing different viral load assays. Shown are comparisons of the Amplicor assay with the TaqMan version 1 (C) or version 2 (D) assay and of the RealTime assay with the TaqMan version 1 (E) or version 2 (F) assay. The bias (mean difference between the assays) is shown as a solid horizontal line, and the 95% limits of agreement are shown with dashed lines at ±1.96 standard deviations.
RESULTS

The correlations between the viral load assays performed on the same sets of samples are shown in Fig. 1A and B. Compared to the Amplicor assay, the TaqMan version 1 assay had a correlation coefficient (Pearson’s R value) of 0.90, and the TaqMan version 2 assay had an R value of 0.97. The correlation coefficients of the TaqMan version 1 and version 2 assays were also 0.90 and 0.97 compared to the RealTime assay. Although there were a small number of data points comparing the Amplicor and RealTime tests (n = 56), these assays had a correlation coefficient of 0.88 (data not shown).

We then restricted our analyses to samples where at least one assay gave a viral load result of <1,000 copies/ml, which is a common definition of low-level viremia. The lower correlation coefficients were obtained when the analyses were performed on low-level viremia samples (Table 1). For example, at <1,000 copies/ml, the RealTime and TaqMan version 2 assays had R values of approximately 0.8 compared with an R value of 0.97 for the comparison over the full range of viral load values. All correlations between the assays were lower when the data were restricted to viral loads of <1,000 copies/ml (Table 1).

Since correlation coefficients do not account for the fact that one assay may provide consistently higher or lower values relative to the other assay, pairwise Bland-Altman plots were used to further assess the level of agreement (Fig. 1C to F). The 95% level of agreement ranged from −0.61 to 0.75 log when the TaqMan version 1 assay was compared with the Amplicor assay (with 55/1,384 [4%] samples outside this range) and ranged from −1.2 to 1.0 log when the RealTime test was compared with the TaqMan version 2 test (with 33/827 [4%] samples outside this range). Samples outside the 95% level of agreement were mostly due to relative underreporting for the TaqMan version 1 test (40 samples and 33 samples for the two comparisons described above, respectively). For the TaqMan version 2 test, the 95% level of agreement versus the Amplicor test was −0.89 to 0.85 (with 1/365 [0.3%] samples below and 14/365 [4%] samples above this range) and versus the RealTime assay was −0.80 to 0.60 (with 19/1,230 [1.5%] samples below and 43/1,230 [3.5%] samples above this range). Comparing the TaqMan version 1 and version 2 assays, approximately 4% (16/415) of the results fell outside the 95% level of agreement, all due to relative underreporting for the TaqMan version 1 test (data not shown).

As previously reported (35), HIV subtype may have partially contributed to assay discordance. The HIV subtype was available for 1,493 of the 4,221 samples (35%). Analyses involving the Amplicor assay were excluded due to a low proportion having HIV subtype information (10% [175/1,749]). Of 25 samples for which the TaqMan version 1 test gave results of at least 1.5 log copies below those of the RealTime test, the most common subtypes were CRF02 (8 samples), subtype F (5 samples), subtype B (4 samples), CRF01 (3 samples), CRF09 (2 samples), CRF14 (2 samples), and subtype G (1 sample). Of 3 samples for which the TaqMan version 2 results were >1.5 log copies below those of the RealTime test, the subtypes were C, CRF01, and CRF02. There were 7 samples that TaqMan version 1 underquantified by >1.5 log compared to TaqMan version 2; 6 of these were CRF02 and 1 was subtype G.

Next, we analyzed assay discordance at a threshold of either 50 or 200 copies/ml (Tables 2 and 3). There was higher concordance between the assays when detectability was defined as ≥200 copies/ml than when defined as ≥50 copies/ml. Overall, 27% (63/230) and 13% (73/569) of the samples with RealTime assay results of <50 copies/ml were above the 50-copies/ml threshold using the TaqMan version 1 and TaqMan version 2 assays, respectively, with median (range) HIV-1 RNA levels of 84 (50 to 394) copies/ml and 99 (51 to 1,620) copies/ml, respectively. For samples with Amplicor assay viral loads of <50 copies/ml, 73% (598/819) of the samples were >50 copies/ml by the TaqMan version 1 assay and 6%...
Proportion of samples with <50 copies/ml by the Amplicor assay as a function of the TaqMan version 1 or version 2 assay. The proportion of samples with Amplicor results of <50 copies/ml decreased at higher TaqMan version 1 and version 2 strata. (A) Proportion of samples with undetectable viral loads by the Amplicor assay (<50 copies/ml) as a function of the viral load reported by the TaqMan version 1 assay. (B) Similar plot as in panel A but as a function of the viral load reported by the TaqMan version 2 assay. For both TaqMan assays, when either had a viral load result of up to 100 copies/ml, a majority of the corresponding samples were actually undetectable by the Amplicor assay.
(10/172) of the samples were >50 copies/ml by the TaqMan version 2 assay, with median (range) HIV-1 RNA levels of 76 (50 to 247) copies/ml and 85 (70 to 164) copies/ml, respectively. Interassay discordance decreased substantially when a 200-copies/ml threshold was used (Table 3). For example, the discordance between the Amplicor and TaqMan version 1 assays dropped from 73% at a 50-copies/ml threshold to 5% at a 200-copies/ml threshold.

Consistent with the findings of higher concordance at higher thresholds, the percentage of results with <50 copies/ml by the Amplicor test decreased in a stepwise pattern as the reported viral load using the TaqMan version 1 test increased (Fig. 2A). The point at which the proportion of Amplicor test results with >50 copies/ml reached a majority of samples occurred in the range of 100 to 124 copies/ml according to the TaqMan version 1 test. In other words, most samples (71%) with TaqMan version 1 test results up to 125 copies/ml were actually <50 copies/ml according to the Amplicor test, with a median of 49 copies/ml (interquartile range [IQR], 49 to 59). For the TaqMan version 2 test, this threshold was approximately 100 copies/ml (Fig. 2B).

**DISCUSSION**

We present here a large comparison of >4,000 paired viral load assay results. We also assessed interassay discordance at low viral load thresholds and found a lower interassay correlation at viral loads of <1,000 copies/ml, compared to the full dynamic ranges of the assays. This range is of clinical importance since many patients are on suppressive or near-suppressive antiretroviral therapies. Indeed, in this study, we deliberately sought a high number of low viral load data points in order to assess the concordance between the assays when performed near their lower limits of quantitation.

We found that correlation between the assays was lower at low viral loads. Discordance decreased when a threshold of 200 copies/ml was used rather than when a 50-copy threshold was used. As was previously reported (19, 36), we confirm that the TaqMan version 2 assay does not appear to have the severe underquantification reported with the TaqMan version 1 assay. However, both TaqMan assays have higher rates of detectability than the Amplicor and RealTime assays. There are a number of factors that may contribute to assay discordance, including viral blips, sample handling, contamination, or differences in the assay primers used (14, 31, 37–42). Furthermore, the assays themselves have inherently lower precision, reproducibility, and sensitivity at the lower ends of their dynamic ranges, which likely contributes to interassay disagreement at low viral loads. Assays exist to quantify viral loads even to single-copy levels, and commercial assays can be modified to accommodate low viral loads (43–46). However, these tests often require large volumes of blood plasma, up to 30 ml in some cases (43), and this precludes their routine use in the clinic setting.

Our data also indicate that a viral load of approximately 100 to 125 copies/ml according to the TaqMan version 1 or TaqMan version 2 assays may correspond to a threshold of detectability of >50 copies/ml by the Amplicor assay, an HIV viral load assay that was widely used for >10 years. Further, we found that interassay concordance for all viral load assays is much higher at a 200-copies/ml threshold than at a 50-copies/ml threshold. Some HIV clinical guidelines (4, 33) have defined higher cutoffs of 200 to 250 copies/ml for low-level viremia, partially in response to the higher rates of detectability seen with this next generation of assays. Thus, the prevalence and extent of low-level viremia may depend on the specific viral load assay in use, and the clinical relevance of the detectable viral load may likewise vary with the assay and threshold defining what is detectable.

The strengths of this data set are its size and the focus on low viral loads that are close to the limits of detection of the assays, but there are also some limitations which should be acknowledged. While the focus of this study was on low viral load levels, there is also variation at the high viral load strata (47). As this was an international collaboration, various methods were used to generate the assay results, and biases inherent in the collections of convenience sample sets in some centers may have had unknown effects. Selection bias may have also influenced the results; for instance, the Vancouver site deliberately retested samples with the Amplicor assay that had TaqMan version 1 assay viral loads of <250 copies/ml (29). The handling procedures varied at different sites as well, and this has been known to impact the results (48), even though all the laboratories reported using EDTA tubes for sample collection and storage. The lack of clinical follow-up for patients experiencing low-level viremia makes the interpretation of these results difficult to situate in a clinical context. Finally, older assays were examined in this study. The TaqMan version 1 assay is being phased out or is no longer in use, and the Amplicor assay has been discontinued. However, although comparisons to the Amplicor assay are of limited relevance to current practice, they are essential to contextualizing newer assays, since the Amplicor assay has dominated the field for so long.

Additional data are needed to determine the impact of low-level viremia on outcomes, such as higher viral rebound, CD4 count, therapy changes, and drug resistance, as well as what the cost impacts are for recalling patients for retesting and changing treatment (41). The high variability around the threshold of detectability of the viral load assays should be noted, since many patients have viral loads in this range. This variability makes defining low-level viremia itself difficult to achieve. We found that agreement between these assays was improved using a 200-copies/ml threshold. Indeed, this threshold is consistent with the current HIV treatment guidelines from the Department of Health and Human Services (DHHS) (4). However, this study does not include clinical follow-up, making firm clinical guidelines difficult to establish based on these data. We suggest that similar large-scale collaborations focusing on low-level viremia that are paired with clinical follow-up would be extremely valuable in establishing such clinical guidelines.

**ACKNOWLEDGMENTS**

We thank Antoinette van Kessel and the European Society for Antiviral Resistance for their assistance in collecting the data set.

Bryan Cobb is an employee of Roche, which manufactures some of the viral load assays used in this study. Roche contributed data comparing the Amplicor and TaqMan version 2 assays.

L.C.S., B.C., A.M.G., P.R.H., M.P., C.S.-D., C.V., M.W., and A.M.J.W. formed the writing committee for this article, and they contributed data and contributed to the analysis and writing of the paper. A.A., J.B., T.B., J.B.H., J.-C.K., S.Z.L., M.M.L., O.M., R.S., J.T., K.V.L., and L.V. formed the scientific committee for this article, and they contributed data and collaborative, various methods were used to generate the assay results, and biases inherent in the collections of convenience sample sets in some centers may have had unknown effects. Selection bias may have also influenced the results; for instance, the Vancouver site deliberately retested samples with the Amplicor assay that had TaqMan version 1 assay viral loads of <250 copies/ml (29). The handling procedures varied at different sites as well, and this has been known to impact the results (48), even though all the laboratories reported using EDTA tubes for sample collection and storage. The lack of clinical follow-up for patients experiencing low-level viremia makes the interpretation of these results difficult to situate in a clinical context. Finally, older assays were examined in this study. The TaqMan version 1 assay is being phased out or is no longer in use, and the Amplicor assay has been discontinued. However, although comparisons to the Amplicor assay are of limited relevance to current practice, they are essential to contextualizing newer assays, since the Amplicor assay has dominated the field for so long.

Additional data are needed to determine the impact of low-level viremia on outcomes, such as higher viral rebound, CD4 count, therapy changes, and drug resistance, as well as what the cost impacts are for recalling patients for retesting and changing treatment (41). The high variability around the threshold of detectability of the viral load assays should be noted, since many patients have viral loads in this range. This variability makes defining low-level viremia itself difficult to achieve. We found that agreement between these assays was improved using a 200-copies/ml threshold. Indeed, this threshold is consistent with the current HIV treatment guidelines from the Department of Health and Human Services (DHHS) (4). However, this study does not include clinical follow-up, making firm clinical guidelines difficult to establish based on these data. We suggest that similar large-scale collaborations focusing on low-level viremia that are paired with clinical follow-up would be extremely valuable in establishing such clinical guidelines.

**REFERENCES**


37. Adachi D, Benedet M, Tang JWT, Gloherty G. 2013. Comparative evaluation of plasma preparation tube (PPT) and EDTA samples using Roche’s Cobas TaqMan HIV-1 and Abbott’s RealTime HIV-1 v. 4.00 assays, poster TUPE262. 7th IAS Conference on HIV Pathogenesis and Treatment, Kuala Lumpur, Malaysia, 30 June to 3 July 2013.

