Comparative performance of HIV-1 RNA load assays at low viral load levels: Results of an international collaboration

The International Viral Load Assay Collaboration

A complete author list as below:

Luke C Swenson1~, Bryan Cobb2, Anna Maria Geretti3, P Richard Harrigan1, Mario Poljak4, Carole Seguin-Devaux5, Chris Verhofstede6, Marc Wirden7, Alessandra Amendola8, Jurg Boni9, Thomas Bourlet10, Jon B. Huder9, Jean-Claude Karasi11, Snjezana Zidovec Lepej12, Maja M. Lunar4, Odette Mukabayire11, Rob Schuurman13, Janez Tomažič4, Kristel Van Laethem14, Linos Vandekerckhove6, Annemarie MJ Wensing13*

~First author, *Corresponding author

Writing Committee:
Contributed data, and to the analysis and writing of the paper
LCS1~, BC2, AMG3, PRH1, MP4, CS-D5, CV6, MW7, AMJW13*

Scientific Committee:
Contributed data and formative input
AA8, JB9, TB10, JBH9, J-CK11, SZL12, MML4, OM11, RS13, JT14, KVL14, LV6

Running head: Viral load assay comparison

Word count: 2112

Abstract: 117

Corresponding Author and requests for reprints:
Annemarie M.J. Wensing, MD, PhD, Department of Virology, Medical Microbiology, University Medical Center Utrecht, G04.614, Heidelbergaan 100, 3584 CX Utrecht, The Netherlands. Tel: +31 88 7556526; fax: +31 88 7555426; e-mail: a.m.j.wensing@umcutrecht.nl
Affiliations:

1 BC Centre for Excellence in HIV/AIDS, Vancouver, Canada
2 Roche Molecular Systems, 4300 Hacienda Drive, Pleasanton, CA 94588
3 Institute of Infection and Global Health, University of Liverpool, UK
4 University of Ljubljana, Institute of Microbiology and Immunology, Ljubljana, Slovenia
5 Laboratory of Retrovirology, CRP-Santé, Luxembourg, Luxembourg
6 AIDS Reference Laboratory, Ghent University, De Pintelaan 185, 9000 Gent, Belgium
7 AP-HP, Pitié-Salpêtrière Hospital, INSERM U 943 and Pierre et Marie Curie University Paris, France
8 Laboratory of Virology, National Institute for Infectious Diseases "Lazzaro Spallanzani", Rome, Italy
9 Institute of Medical Virology, Swiss National Center for Retroviruses, University of Zürich, Zürich, Switzerland
10 Laboratory of Virology, CHU de Saint-Etienne, Saint-Etienne, France
11 Ministry of Health, National Reference Laboratory, Kigali, Rwanda
12 University Hospital for Infectious Diseases, Mirogojska 8, Zagreb, Croatia
13 University Medical Center Utrecht, Utrecht, Netherlands
14 Clinical and Epidemiological Virology, Rega Institute for Medical Research, KU Leuven, Leuven, Belgium
Abstract

Low-level viraemia during antiretroviral therapy and its accurate measurement have become increasingly relevant. We present an international collaboration of 4,221 paired plasma viral load (pVL) results from four commercial assays, emphasizing data with low pVL.

The assays compared were the Abbott RealTime assay and the Roche Amplicor, TaqMan v1 and v2 assays. Correlation between assays was 0.90-0.97. However, at low pVL, correlation fell to 0.45-0.85. Higher inter-assay concordance was observed when detectability was defined as 200-copies/mL versus 50-copies/mL.

A pVL of ~100-125 copies/mL by TaqMan v1/v2 corresponded best to a 50-copies/mL threshold by Amplicor.

Correlation and concordance between viral load assays were lower at low pVL.

Clear guidelines are needed on the clinical significance of low-level viraemia.

Key words: HIV; viral load assays; Amplicor; TaqMan; RealTime; low-level viraemia
Background

The level of HIV-1 RNA in 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 (version 1 and version 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 use an “undetectable” viral load by Amplicor (below 50 HIV-1 RNA copies/mL (3)) as the end-point for measuring efficacy of antiretroviral therapy, and guidelines have adopted the same threshold as the goal of therapy for all patients, regardless of previous treatment experience (1, 4–6). After many years of Amplicor use, a real-time PCR-based, Roche COBAS® AmpliPrep/ COBAS® TaqMan® Real-Time Quantitative Human Immunodeficiency Virus Type HIV-1 Test (“TaqMan”) began replacing Amplicor in early 2008. There are two versions of TaqMan; version 2 has now replaced version 1 across Europe, Asia and North America. TaqMan versions 1 and 2 had respective lower limits of quantitation of 40 and 20 copies/mL (7–9). Other assays based on real-time PCR technology have also been implemented in clinical settings, including the Abbott RealTime HIV-1 Assay (“RealTime”) and the QIAGEN artus HIV-1 QS-RGQ Assay (10, 11), which have lower limits of quantitation of 40 and 45 copies/mL (12, 13).

Although generally these assays tend to give roughly comparable results (14), their manufacturers have reported that variation and error tends to increase at their lower limits of quantitation (12, 15). Furthermore, there have been reports that real-time PCR-based assays tend to identify more patients as having low-
level viraemia than Amplicor (14, 16–20). Multiple studies have documented variation amongst these assays when measuring at low viral load levels (9, 18, 20–26)

Therapy outcomes have been shown to be influenced by low-level viraemia in some studies (21, 27–33) but not others (34–37), in part reflecting differences in measured outcomes. However, the overall clinical relevance of low-level viraemia remains controversial. Accordingly, some clinical guidelines (5, 38) 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 viraemia be reviewed due to a potential risk of virological rebound.

A significant clinical question has been what the relevance of low-level viraemia is. In order to properly address this question, an assessment of the fundamental agreement of various assays for defining low-level viraemia must be undertaken. Given that the above viral load assays differ in their reported lower limits of quantitation and detection (LLQ & LLD) (39), and given the lack of definitive guidelines on the management of low-level viraemia, 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 v1 and TaqMan v2 viral load assays at low viral loads.
Methods

A total of 14 sites from Europe, North America, and Africa contributed 4,221 paired viral load results obtained by different assays performed from the same sample. The number of data-points per comparison were as follows: Amplicor vs. TaqMan v1 (N=1,384); Amplicor vs. TaqMan v2 (N=365); RealTime vs. TaqMan v1 (N=827); RealTime vs. TaqMan v2 (N=1,230); and TaqMan v1 vs. TaqMan v2 (N=415). Subtype data, where provided, was collected independently at each site by various in-house methods. Ethical approval at individual collaborating centers was based on local regulations. There were insufficient numbers of paired results from both Amplicor and RealTime, so comparisons between them were excluded. Correlation and Bland-Altman analyses were performed using paired data points. If a viral load value fell below the LLQ of a given assay, or was classified as not detected, the value was imputed as 1 copy less than the LLQ (e.g. 49 for a LLQ of 50 copies/mL).
Results

The correlation between viral load assays performed on the same sets of samples are shown in Figures 1A and 1B. Compared to Amplicor, TaqMan v1 had a correlation coefficient (Pearson’s R value) of 0.90; and TaqMan v2 had an R value of 0.97. The correlation coefficients of TaqMan v1 and v2 were also 0.90 and 0.97 when compared to RealTime. Although there were a small number of data-points comparing Amplicor and RealTime (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 <1,000 copies/mL, a common definition of low-level viraemia. Lower correlation coefficients were obtained when analyses were performed on low-level viraemia samples (Table 2). For example, at <1000 copies, RealTime and TaqMan v2 had R values of approximately 0.8 compared with R=0.97 for the comparison over the full range of viral load values. All correlations between assays were lower when the data were restricted to viral loads <1,000 copies/mL (Table 2).

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 (Figures 1C-1F). The 95% level of agreement ranged from -0.61 to 0.75 log when TaqMan v1 was compared with Amplicor (with 55/1384 (4%) samples outside this range) and ranged from -1.2 to 1.0 log when RealTime was compared with Amplicor (33/827 (4%) samples outside). Samples outside the 95% level of
agreement were mostly due to relative under-reporting by TaqMan v1 (40 samples and 33 samples, respectively). For TaqMan v2, the 95% level of agreement versus Amplicor was $-0.89$ to $0.85$ (with $1/365$ (0.3%) below and $14/365$ (4%) samples above) and versus RealTime was $-0.80$ to $0.60$ ($19/1230$ (1.5%) below and $43/1230$ (3.5%) above). Comparing TaqMan v1 with TaqMan v2, approximately $4\%$ (16/415) of results fell outside the 95% level of agreement, all due to relative under-reporting by TaqMan v1 (data not shown).

As previously reported (40), HIV subtype may have partially contributed to assay discordance. HIV subtype was available for 1493 of 4221 samples (35%). Analyses involving Amplicor were excluded due to a low proportion having HIV subtype information (10%, 175/1749). Of 25 samples where TaqMan v1 gave results at least 1.5 log copies below RealTime, the most common subtypes were CRF02 (8 samples), subtype F (5 samples), subtype B (4 samples), CRF01 (3 samples), CRF09 and CRF14 (2 samples each), and subtype G (1 sample). Of 3 samples where TaqMan v2 results were $>1.5$ log copies below RealTime, the subtypes were C, CRF01, and CRF02. There were 7 samples that TaqMan v1 underquantified by $>1.5$ log compared to TaqMan v2; 6 of these were CRF02 and 1 was subtype G.

Next, we analyzed assay discordance at thresholds of either 50 or 200 copies/mL (Tables 1A and 1B). There was higher concordance between assays when “detectability” was defined as $\geq 200$ copies versus $\geq 50$ copies. Overall, 27% (63/230) and 13% (73/569) of samples with RealTime results $<50$ copies/ml were above the 50 copies/ml threshold by Taqman v1 and Taqman v2,
respectively, with median (range) HIV-1 RNA levels of 84 (50–394) and 99 (51-1,620) copies/ml. For samples with Amplicor viral loads <50 copies/ml, 73% (598/819) of samples were >50 copies/ml by TaqMan v1 and 6% (10/172) by TaqMan v2, with median (range) HIV-1 RNA levels of 76 (50-247) and 85 (70-164), respectively. Inter-assay discordance decreased substantially when a 200 copies/mL threshold was used (Table 1B). For example, the discordance between Amplicor and TaqMan v1 dropped from 73% a 50 copies/ml threshold to 5% at a 200 copies/ml threshold.

Consistent with higher concordance at higher thresholds, the percentage of results <50 copies/mL by Amplicor decreased in a step-wise pattern as the reported viral load by TaqMan v1 increased (Figure 2A). The point at which the proportion of Amplicor results >50 copies reached a majority of samples occurred in the range of 100-124 copies/mL according to TaqMan v1. In other words, most samples (71%) with TaqMan v1 results up to 125 copies/mL were actually <50 copies/ml by Amplicor, with a median of 49 copies/mL (IQR: 49-59). For TaqMan v2, this threshold was approximately 100 copies/mL (Figure 2B).
Discussion

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

We find that correlation between assays was lower at low viral loads. Discordance decreased when a threshold of 200 copies/mL was used rather than a 50 copy threshold. As previously reported (24, 41), we confirm the TaqMan v2 assay does not appear to have the severe under-quantification reported with TaqMan v1. However, both TaqMan assays have higher rates of detectability than Amplicor and RealTime. There are a number of factors that may contribute to assay discordance, including viral blips, sample handling, contamination, or differences in assay primers (19, 36, 42–47). Furthermore, the assays themselves have inherently lower precision, reproducibility, and sensitivity at the lower ends of their dynamic ranges which likely contributes to inter-assay disagreement at low viral loads (3, 7, 12). Assays exist to quantify viral loads to even single-copy levels, and commercial assays can be modified to accommodate low viral loads (48–51). However, these often require large
volumes of blood plasma – up to 30 milliliters in some cases (48) – and this precludes their routine use in the clinic.

Our data also indicate that a viral load of approximately 100-125 by TaqMan v1 or TaqMan v2 may correspond to a threshold of detectability >50 copies/ml by Amplicor, an HIV viral load assay that was widely used for over 10 years. Further, we found that inter-assay concordance for all viral load assays is much higher at a 200 copies/ml threshold compared to a 50 copies/ml threshold. Some HIV clinical guidelines (5, 38) have defined higher cutoffs of 200-250 partially in response to the higher rates of detectability seen with this next generation of assays. Thus the prevalence and extent of low-level viraemia may depend a great deal on the specific viral load assay in use, and the clinical relevance of “detectable” viral load may likewise vary with the assay and threshold defining what is detectable.

Strengths of this dataset are its size and the focus on low viral loads close to the assays’ limits of detection, 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 high viral load strata (52). As this was an international collaboration, varying methods were used to generate the assay results, and biases inherent in collections of convenience sample sets in some centres may have had unknown effects. Selection bias may have also influenced results; for instance, the Vancouver site deliberately retested samples with Amplicor that had TaqMan v1 viral loads <250 copies/mL (34). Handling procedures varied at different sites as well, and this has been known to impact results (53), though all
labs reported using EDTA tubes for sample collection and storage. The lack of clinical follow-up for patients experiencing low-level viraemia makes the interpretation of these results difficult to situate in a clinical context. Finally, older assays were examined in this study. TaqMan v1 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 Amplicor has dominated the field for so long.

Additional data is needed to determine the impact of low-level viraemia on outcomes like higher viral rebound, CD4 count, therapy changes, and drug resistance; and also what the cost impacts are for recalling patients for retesting and changing treatment (46). The high variability around the threshold of detectability of viral load assays should be noted, since many patients have viral loads in this range. This variability makes defining low-level viraemia itself difficult to achieve. We found agreement amongst these assays was improved using a 200 copy threshold. Indeed, this threshold is consistent with the current HIV treatment guidelines of the DHHS (5). 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 focused on low-level viraemia and paired with clinical follow-up would be extremely valuable in establishing such clinical guidelines.
References


3. Roche Molecular Systems Inc. AMPLICOR HIV-1 MONITOR TEST. FDA.gov.


7. Roche Molecular Systems Inc. COBAS® AmpliPrep/COBAS® TaqMan® HIV-1 Test.

8. Roche Molecular Systems Inc. COBAS® AmpliPrep/COBAS® TaqMan® HIV-1 Test, v2.0.


Consequences at the Cohort Level. PLoS One 8:e74024.


Acknowledgments

The authors would like to thank Antoinet van Kessel and the European Society for Antiviral Resistance for assistance in collecting the dataset. 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 v2 assays.
Figure 1: Comparisons amongst viral load assays

The upper two panels show the correlation of both TaqMan assays with either the Amplicor assay (Figure 1A) or the RealTime assay (Figure 1B). The lower four panels show Bland-Altman difference plots comparing different viral load assays. For Figures 1A and 1B, comparisons involving TaqMan v1 are shown with filled grey circles, while those involving TaqMan v2 are shown with open squares. The R² values for all correlations are shown in the legends. Figures 1C and 1D show comparisons of Amplicor with TaqMan v1 or v2, respectively. Figures 1E and 1F show comparisons of RealTime with TaqMan v1 or v2, respectively. The bias (mean difference between assays) is shown as the solid horizontal line, and the 95% limits of agreement are shown with dashed lines at ±1.96 standard deviations.

Figure 2: The proportion of samples below 50 copies/mL by Amplicor as a function of TaqMan v1 or v2

The proportion of samples with Amplicor results <50 copies/mL decreased at higher TaqMan v1 and v2 strata. Figure 2A shows the proportion of samples with undetectable viral loads by Amplicor (<50 copies/mL) as a function of the viral load reported by TaqMan v1. Figure 2B shows a similar plot but as a function of the viral load reported by TaqMan v2. For both TaqMan assays, when either gave a viral load result of up to 100 copies/mL a majority of corresponding samples were actually undetectable by Amplicor.
Table 1: Discordance for HIV-1 RNA levels below 50 or 200 copies/mL

The tables show the percentage of results that were discordant between assays at a threshold of 50 copies/mL (1A) or 200 copies/mL (1B). The percentage of discordant samples decreased for all comparisons when a threshold of 200 copies/mL was used, compared to a 50 copy threshold.

Table 2: Correlation coefficients for inter-assay comparisons at viral loads <1000 copies/mL

The correlation coefficients of inter-assay comparisons were calculated on results with restriction to where at least one viral load assay gave a result <1000 copies/mL. The assay restricted to <1000 is shown in the lefthand column, and its comparator assay is indicated in the upper row. For example, When TaqMan v1 was restricted to values <1000, matching Amplicor results had an R value of 0.45. Conversely, when Amplicor was restricted to <1000, matching TaqMan v1 results had an R value of 0.52.
Table 1A: Discordance for HIV-1 RNA levels below 50 copies/mL

<table>
<thead>
<tr>
<th>Discordant results ≥50</th>
<th>Amplicor &lt;50</th>
<th>RealTime &lt;50</th>
<th>TaqMan v1 &lt;50</th>
<th>TaqMan v2 &lt;50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplicor &lt;50</td>
<td>-</td>
<td>NA</td>
<td>73% (598/819)</td>
<td>6% (10/172)</td>
</tr>
<tr>
<td>RealTime &lt;50</td>
<td>NA</td>
<td>-</td>
<td>27% (63/230)</td>
<td>13% (73/569)</td>
</tr>
<tr>
<td>TaqMan v1 &lt;50</td>
<td>11% (27/248)</td>
<td>23% (49/216)</td>
<td>-</td>
<td>24% (50/206)</td>
</tr>
<tr>
<td>TaqMan v2 &lt;50</td>
<td>5% (8/170)</td>
<td>7% (40/536)</td>
<td>25% (51/207)</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1B: Discordance for HIV-1 RNA levels below 200 copies/mL

<table>
<thead>
<tr>
<th>Discordant results ≥200</th>
<th>Amplicor &lt;200</th>
<th>RealTime &lt;200</th>
<th>TaqMan v1 &lt;200</th>
<th>TaqMan v2 &lt;200</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplicor &lt;200</td>
<td>-</td>
<td>NA</td>
<td>5% (58/1161)</td>
<td>3% (7/204)</td>
</tr>
<tr>
<td>RealTime &lt;200</td>
<td>NA</td>
<td>-</td>
<td>10% (37/356)</td>
<td>7% (49/728)</td>
</tr>
<tr>
<td>TaqMan v1 &lt;200</td>
<td>7% (85/1188)</td>
<td>16% (59/378)</td>
<td>-</td>
<td>7% (25/336)</td>
</tr>
<tr>
<td>TaqMan v2 &lt;200</td>
<td>3% (7/204)</td>
<td>3% (22/701)</td>
<td>1% (3/314)</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 2: Correlation coefficients for inter-assay comparisons at viral loads <1000 copies/mL

<table>
<thead>
<tr>
<th></th>
<th>Amplicor &lt;1000</th>
<th>RealTime &lt;1000</th>
<th>TaqMan v1 &lt;1000</th>
<th>TaqMan v2 &lt;1000</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Amplicor</strong></td>
<td>-</td>
<td>NA</td>
<td>0.52</td>
<td>0.85</td>
</tr>
<tr>
<td><strong>RealTime</strong></td>
<td>NA</td>
<td>-</td>
<td>0.74</td>
<td>0.82</td>
</tr>
<tr>
<td><strong>TaqMan v1</strong></td>
<td>0.45</td>
<td>0.54</td>
<td>-</td>
<td>0.69</td>
</tr>
<tr>
<td><strong>TaqMan v2</strong></td>
<td>0.80</td>
<td>0.78</td>
<td>0.68</td>
<td>-</td>
</tr>
</tbody>
</table>