Precision across the Analytical Measuring Range of a Quantitative Real-Time PCR Assay for Cytomegalovirus Detection among Three Clinical Laboratories[▽]

Thomas E. Grys,^{1,2*} Doreen L. Duquette,¹ Bruce White,³ Cole Irish,²
D. Jane Hata,³ and Bobbi S. Pritt²

Molecular Microbiology Laboratory, Mayo Medical Laboratories New England, Andover, Massachusetts¹; Division of Clinical Microbiology, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota²; and Division of Clinical Microbiology, Department of Laboratory Medicine and Pathology, Mayo Clinic, Jacksonville, Florida³

Received 9 March 2011/Returned for modification 3 May 2011/Accepted 26 May 2011

This study measured the precision of a quantitative laboratory-developed real-time PCR test for cytomegalovirus performed at three different clinical laboratories that use the same methodology. The overall standard deviation (adjusted for analyte level) was 0.18 log₁₀ copies/ml, and there was no significant relationship between standard deviation and analytical measuring range.

Previous studies have shown that laboratory-developed viral load (VL) assays differ dramatically among various instruments, reagents, and calibration schemes (5, 9, 10, 12). For this reason, it is strongly recommended that individual patient VLs should always be tracked using the same methodology at the same laboratory. Furthermore, it is important to understand the precision of the assay to clarify when a difference between two subsequent measurements of the VL represents a significant change or when the difference is insignificant due to the inherent imprecision of the assay.

The implementation of international cytomegalovirus (CMV) reference materials (4, 7) will aid in the harmonization of different CMV test methods, but the question of within-assay precision will remain. Therefore, we sought to quantify the precision of a quantitative real-time PCR assay across three clinical laboratories that use standardized procedures, analyte-specific reagents (ASRs), instruments, calibrators, and specimen requirements. Second, we analyzed the data to determine whether precision varied over the analytical measuring range (AMR). Finally, we determined repeatability (within-challenge precision for each site) and reproducibility (within-challenge precision among all sites) of the assay at three analyte levels over time.

Over a 6-month period, four challenges of blinded split samples ($n = 180$ in total) were sent to three testing sites within Mayo Medical Laboratories and tested as a part of routine diagnostic runs. Sample group 1 (SG1) was composed of whole virus suspended in a commutable EDTA plasma matrix (AcroMetrix, part of Life Technologies, Benicia, CA). Sample group 2 (SG2) was normal human EDTA plasma spiked with a high CMV VL patient specimen.

Challenges were sent on days 0, 14, 99, and 178 (Table 1). Challenges sent to each site on days 0 and 14 contained only

samples from SG1, comprising four CMV levels (0, 3.7, 4.7, and 5.7 log₁₀ copies/ml), each level in triplicate for each site. Each replicate was tested once, yielding 9 data points for each level on each day. The day 99 challenge was the same for days 0 and 14 but also included SG2 material at 5.8 log₁₀ copies/ml, also in triplicate for each site. The day 178 challenge was the same as day 99 but included two additional levels of SG2 (4.2 and 4.5 log₁₀ copies/ml), each in triplicate for each site. Overall, 180 samples were tested, including 36 data points for each level of SG1 ($n = 144$) and 9 or 18 points for each level of SG2 ($n = 36$).

The assay was a laboratory-developed test using LightCycler CMV UL54 primer/hybridization probes (ASRs) (Roche Diagnostics, Indianapolis, IN) as described by Razonable et al (11), targeting the polymerase gene. Briefly, extraction of the samples (0.2 ml input) was performed using a MagNA Pure LC total nucleic acid isolation kit on a MagNA Pure LC instrument (Roche Diagnostics, Indianapolis, IN) with 0.1 ml elution. The reaction volume was composed of 15 μ l master mix (1 $\times$ Roche FastStart DNA Master HybProbe) and 5 μ l sample extract with 45 cycles of PCR using a LightCycler 2.0 instrument (Roche Diagnostics). The assay was calibrated using the CMV UL54 template set (Roche Diagnostics). The analytical measuring range was 3.3 to 8.3 log₁₀ copies/ml.

All results were qualitatively correct, and 96% of specimens were within 0.50 log₁₀ copies/ml of the overall mean value of each respective analyte level. To determine the overall precision, data from SG1 samples was used because this material was sent with each of the four challenges. The consistency of the materials, number of replicates, and time frame of testing make for a robust data set for statistical analysis. Standard deviations (SDs) adjusted for differences between analyte group levels were estimated with the square root of the mean squared error (RMSE) from analysis-of-variance (ANOVA) models calculated using JMP software (2008; SAS Institute Inc., Cary, NC), and the overall precision in terms of SD was 0.18 log₁₀ copies/ml (Table 1).

To investigate the relationship of precision across the AMR, the second sample set, SG2, was also included. Interestingly,

* Corresponding author. Mailing address: Molecular Microbiology Laboratory, Mayo Medical Laboratories New England, 160 Dascomb Rd., Andover, MA 01810. Phone: (978) 253-6187. Fax: (978) 658-0439. E-mail: gryst.thomas@mayo.edu.

[▽] Published ahead of print on 8 June 2011.

TABLE 1. Challenge logistics and precision at each level

Sample group	No. of samples per challenge at indicated level (log ₁₀ copies/ml) ^{a,b}							Total no. of samples per challenge
	SG1 ^c				SG2			
	0	3.7	4.7	5.7	4.2	4.5	5.8	
Day 0 challenge	9	9	9	9	0	0	0	36
Day 14 challenge	9	9	9	9	0	0	0	36
Day 99 challenge	9	9	9	9	0	0	9	45
Day 178 challenge	9	9	9	9	9	9	9	63
Total no.	36	36	36	36	9	9	18	180

^a Overall SD were as follows: for SG1, 3.7- $\log_{10}$ copies/ml level, 0.22; 4.7 level, 0.18; 5.7 level, 0.15; for SG2, 4.2 level, 0.10; 4.5 level, 0.12; 5.8 level, 0.26 (this last analysis excludes three outlying data points from the SG2 5.8 level, which were all from one site on one challenge date; including them, the SD = 0.39).

^b In the Brown-Forsythe test (ANOVA) for difference of SD between analyte levels for SG1 only, $P = 0.128$. In the Brown-Forsythe test for SG1 and SG2 combined, $P = 0.133$. This analysis excludes three outlying data points from the SG2 5.8 level, which were all from one site on one challenge date.

^c Total RMSE for SG1 was 0.18. MSE (mean squared error) estimates variance after accounting for the different levels used in the challenge. RMSE (root mean square error; the square root of the MSE) is the overall SD adjusted for challenge level.

the SG1 and SG2 data had opposite trends in terms of precision versus AMR (Fig. 1). Part of the discrepancy might be due to the materials used (SG1 samples were a commercial product; SG2 samples were spiked specimens). The highest level (5.8 $\log_{10}$ copies/ml) of the SG2 data had three potential outliers, all from one challenge at one site. Including those three points with all data from SG1 and SG2, the Brown-Forsythe test of equivalence (1) indicated that at least one of the SDs was not equal ($P = 0.001$), with the outlying SD being at the 5.8 $\log_{10}$ copies/ml level, as expected. Omitting these three points changed the Brown-Forsythe test to nonsignificance ($P = 0.133$) (Table 1). Our interpretation was that these three points were outliers, since their consideration has such a strong bearing on the overall analysis. Omitting all results from the SG2 5.8 $\log_{10}$ copies/ml level (and thus avoiding the question of outliers) yielded a result of no difference by the Brown-Forsythe test ($P = 0.091$). Thus, the results from SG1 and SG2 demonstrated that the precision of the assay was not different among the different analyte levels in the AMR.

The relationship of precision over the AMR has implications for the reporting of results if a laboratory wants to include a confidence interval. Since the standard deviation did not demonstrate a significant difference over the AMR, the same confidence interval may be applied to any result that is reported

within the AMR. In our laboratories, we report both a log and an integer value. We also include a 95% confidence interval to help clinicians determine if an observed change in VL between consecutive measurements exceeds the inherent variability of the assay.

Repeatability (within-site precision) and reproducibility (across-site precision) were tracked over the four challenges for the SG1 samples. Repeatability (a pooled SD calculated as the square root of the average of the site-specific variances, for each day at each level) was generally less than 0.15 $\log_{10}$ copies/ml. Reproducibility (calculated as the SD of all results from all sites, for each level on each day) was generally less than 0.25 $\log_{10}$ copies/ml.

One advantage of SG1 was that it was commercially prepared material, prepared with good manufacturing practices, and was purported to be a commutable material (i.e., it should behave the same as a patient specimen and give similar results by different methods that share a calibration). The use of commutable calibrators has been shown to increase agreement between laboratories (2). The disadvantage of SG1 was that it may not always accurately represent patient specimens. The SG2 data similarly suffered from the fact that although the samples were derived from positive patient material, they remain a constructed sample set, and there were fewer data points for the analyte levels in SG2. Also, our study did not address biological variation (i.e., variability of VL due to host factors). In HIV-1 infections, the biological variability is thought to contribute about 0.3 $\log_{10}$ copies/ml to the 0.5- $\log_{10}$ copies/ml threshold that is considered the minimum significant change (3, 6, 8).

This study design may demonstrate a best-case scenario for multisite comparability of a quantitative laboratory-developed assay. The primary study outcome provides a standard deviation of 0.18 $\log_{10}$ copies/ml as a point of reference for the standardization of laboratory-developed tests among multiple sites. This should be helpful as laboratories move to standardize calibration and reporting units and set acceptance criteria based on the World Health Organization International Standard for CMV and a forthcoming CMV standard reference material from the National Institute of Standards and Technology (4, 7).

FIG. 1. Scatter plot of standard deviation versus analytical measuring range for SG1 and SG2 data. If one run of apparent outliers at level 5.8 (SG2) is excluded, the SD changes from 0.39 to 0.26, indicated by the asterisk.

We thank Acrometrix for providing the materials in SG1. We are indebted to Sarah Jenkins of the Division of Biomedical Statistics & Informatics, Mayo Clinic, Rochester, MN, for her input on the study design and analysis and for her thoughtful review of the manuscript.

REFERENCES

1. **Brown, M. B., and A. B. Forsythe.** 1974. Robust tests for equality of variances. *J. Am. Stat. Assoc.* **69**:364–367.
2. **Caliendo, A. M., et al.** 2009. A commutable cytomegalovirus calibrator is required to improve the agreement of viral load values between laboratories. *Clin. Chem.* **55**:1701–1710.
3. **CLSI.** 2010. Quantitative molecular methods for infectious diseases; approved guideline, 2nd ed. CLSI document MM06-A2. Clinical and Laboratory Standards Institute, Wayne, PA.
4. **Fryer, J. F., H. A. R. Anderson, P. D. Minor, and the Collaborative Study Group.** 2010. Collaborative study to evaluate the proposed 1st WHO International Standard for human cytomegalovirus (HCMV) for nucleic acid amplification (NAT)-based assays. WHO/BS/10.2138. WHO ECBS report. WHO, Geneva, Switzerland.
5. **Hayden, R. T., et al.** 2008. Multicenter comparison of different real-time PCR assays for quantitative detection of Epstein-Barr virus. *J. Clin. Microbiol.* **46**:157–163.
6. **Kellogg, J. A., P. V. Atria, J. C. Sanders, and M. E. Eyster.** 2001. Intra- and interlaboratory variabilities of results obtained with the Quantiplex human immunodeficiency virus type 1 RNA bDNA assay, version 3.0. *Clin. Diagn. Lab. Immunol.* **8**:560–563.
7. **National Institutes of Standards and Technology.** 5 October 2010, posting date. Certification of a standard reference material for cytomegalovirus. 2011. National Institute of Standards and Technology, Gaithersburg, MD.
8. **Panel on Antiretroviral Guidelines for Adults and Adolescents.** 2011. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents, p. 9–10. Department of Health and Human services, Washington, DC. <http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf>. Accessed 10 January 2011.
9. **Pang, X. L., et al.** 2009. Interlaboratory comparison of cytomegalovirus viral load assays. *Am. J. Transplant.* **9**:258–268.
10. **Preiksaitis, J. K., et al.** 2009. Interlaboratory comparison of Epstein-Barr virus viral load assays. *Am. J. Transplant.* **9**:269–279.
11. **Razonable, R. R., et al.** 2001. Comparative quantitation of cytomegalovirus (CMV) DNA in solid organ transplant recipients with CMV infection by using two high-throughput automated systems. *J. Clin. Microbiol.* **39**:4472–4476.
12. **Wolff, D. J., D. L. Heaney, P. D. Neuwald, K. A. Stellrecht, and R. D. Press.** 2009. Multi-site PCR-based CMV viral load assessment-assays demonstrate linearity and precision, but lack numeric standardization: a report of the Association for Molecular Pathology. *J. Mol. Diagn.* **11**:87–92.