Adding value to antiretroviral proficiency testing

Robin DiFrancesco[*,1], **Charlene R. Taylor**[1], **Susan L. Rosenkranz**[2], **Kelly M. Tooley**[1], **Poonam G. Pande**[3], **Suzanne M. Siminski**[2], **Richard W. Jenny**[4], and **Gene D. Morse**[1]

[1]HIV Clinical Pharmacology Research Program, Translational Pharmacology Research Core, New York State Center of Excellence in Bioinformatics & Life Sciences; School of Pharmacy & Pharmaceutical Sciences, University at Buffalo, Buffalo, NY USA

[2]Frontier Science Technology & Research Foundation, Amherst, NY USA

[3]RTI International, Research Triangle Park, NC USA

[4]New York State Department of Health, Albany, NY USA

Abstract

Background—Clinical trial specimens tested for antiretroviral (ARV) concentrations often require compliance with Clinical Laboratory Improvement Act and/or the Food and Drug Administration bioanalytical guidance.

Experimental—The Clinical Pharmacology Quality Assurance Program (CPQA) designed 8 proficiency testing (PT) rounds over 4 years to assess precision, specificity and stability.

Results—Ten laboratories provided blinded proficiency data to support continued acceptable precision of ARV methods. Specificity samples identified little bias for individual methods; hemolyzed (87%) and lipemic (86%) results were 10% of their control results. Stability was established for ARVs in plasma at −70°C for 2.5–3.6 years.

Conclusion—PT provided by the CPQA assured continued acceptability of individual laboratory assay performances for precision and specificity, and obtained ARV stability during long term storage.

Background

Proficiency testing (PT) in therapeutic drug monitoring (TDM) provides assurance of laboratory accuracy and measures the variability of results between participating laboratories. In the United States, the Clinical Laboratory Improvement Act (CLIA) requires proficiency testing for TDM [1]. The Clinical Pharmacology Quality Assurance Program

(CPQA) provides PT for antiretrovirals (ARVs) in clinical trial protocols conducted by the National Institute of Allergy and Infectious Diseases (NIAID). The CPQA has recently reported data from the initial four PT events [2]. The ARVs tested included 9 HIV-1 protease inhibitors (PI), 3 HIV-1 non-nucleoside reverse transcriptase inhibitors (NNRTI), 7 nucleoside reverse transcriptase inhibitors (NRTI), 1 CCR5 antagonist and 1 integrase strand transfer inhibitor (INSTI). Ten Clinical Pharmacology Laboratories (CPL) participated in the PT program. As shown by these results some ARV are tested more often by CPLs than other ARVs. Prior reports have described the accuracy and variability of ARV concentration measurements across laboratories participating in ARV PT programs where the bioanalytical methods that have been employed are almost exclusively chromatographic methods [3–7]. These reports have also investigated factors that could potentially affect accuracy such as laboratory, sample preparation method, detection method, concentration level, and specific ARV. Stated outcomes have sometimes differed, but these differences have been discussed within the reports.

To validate chromatographic methods used to measure drugs and their metabolites, the US Food and Drug Administration (FDA) or European Medicines Agency (EMA) bioanalytical method validation requirements are most often used [8,9]. Several requirements established during validation are: assay precision and accuracy at three or more ARV concentrations (encompassing the quantitative curve range) as well as the lower limit of quantitation, stability of the ARV in the matrix under various relevant conditions, the absence of sample-matrix effects and identification of any specific interferences (endogenous or concomitant medication). The CPQA program guidelines recommends use of these country guidelines during validation and conducts peer review of its participating laboratories analytical validation reports and resultant standard operating procedures. Many of the networks supported by the CPQA programs require laboratory methods to be approved by this program prior to application to network clinical trial specimens.

While laboratories actively analyzing clinical research samples for common ARV and metabolites may validate chromatographic assays using FDA and EMA guidance, these assays, once implemented, often remain continually active for years afterward. During this extended period of time, although laboratories may follow the required standard acceptance criteria, some may not reevaluate the intra- and inter-assay precision of their method until some form of remediation is needed. Furthermore, multiple laboratories may be testing clinical trial samples to quantitate the same ARV and data could be combined for future analysis without regard to the differences in continued laboratory assay performance. Thus having some measure of continued precision and specificity is of potential valuable.

Recent guidance from the EMA as well as white papers [10] compel testing laboratories to also determine the effects of hemolysis and hyperlipidemia on the accuracy of drug measurements during method development and validation. For example, hemolysis releases hemoglobin into the plasma sample where it may cross-react with an analyte of interest lowering the true concentration or its presence may cause degradation of the analyte [11]. Likewise, lipemia causes volume displacement in the plasma thus decreasing the actual content of plasma in a sample [12]. Both conditions have been shown to cause matrix effects. Examples of reported adverse effects of hemolysis and lipemia using

chromatographic assays coupled to mass spectrometry detection include the degradation of compounds measured in patients receiving treatment for malaria and suppressive matrix effects on the measurements of both drug (olanzapine) and metabolite (desmethyl olanzapine) as well as their isotopic internal standards, respectively [13,14]. The suppression of olanzapine and its metabolite was also reported in hemolyzed samples by Hughes, *et al.*, along with three other assays affected by hemolysis [15]. Older methods validated prior to these initiatives may require retesting for specificity. Furthermore, proficiency testing programs should consider mechanisms to assure continued specificity of laboratory measurements.

Long term analyte stability measured in bioanalytical laboratories during the premarketing phase of a drug tends to be limited to 1–2 years of storage. Post-marketing clinical trials can often take longer to accrue the participants and the participants may be studied for a longer duration. In most cases, especially in international studies, samples will be stored up to the end of the trial for bulk shipment with subsequent bioanalysis. As a result, stability must be extended to include longer periods of time.

In 2000, it was recommended that, in order to assist laboratories in remediation and quality improvement, PT programs include information such as the magnitude of error to laboratories that demonstrate unsatisfactory performance [16]. To provide "value added" efficiency and economy during the analytical phase, the CPQA incorporated samples into the PT panels to allow monitoring of laboratory-specific intra-assay variability, inter-assay variability and method specificity with respect to endogenous factors such as hemolysis and lipemia. The intra- and inter-assay precision testing sought to assess the continued acceptability of ARV methods used to measure ARV in clinical trial samples, albeit within the limitations of the PT program criteria. The precision testing was specifically for reassurance due to the longevity of the ARV methods used by participating laboratories. Specificity testing was supplemental to the PT process and sought to assist laboratories. PT samples from earlier rounds were also included to assess long-term ARV stability in plasma at −70° C using results for three or more laboratories. These measurements occurred during the first eight testing events. While numerous ARVs were included in the testing panels, this report is focused on ARVS of current interest to HIV researchers.

Experimental

Supplies & reagents

Normal plasma (K2ETDA) used for proficiency test samples was purchased (frozen) from Equitech (Kerriville, TX) for rounds 23–27 and Bioreclamation (Westbury, NY, USA) for rounds 28–30. Hemolyzed plasma and lipemic plasma was also purchased from the latter vendor. Normal plasma was filtered using a Whatman 0.2 micron cellulose acetate membrane with a 0.5 micron prefilter (ZAPCAP™; GE Healthcare Life Sciences, Piscataway, NJ, USA). Filtration was necessary to remove insoluble materials that are a direct result of the freeze-thaw process and provide transparent plasma similar to clinical samples. The hemolyzed plasma was graded at 140mg/dl of hemoglobin (versus a normal value 0, slight 100, moderate 200, and gross 400). The lipemic plasma was tested by Clinical Reference Lab using the Cobas Serum Index Gen2 assay (Lenexa, TX, USA) and graded as

severely lipemic (graded 140 versus a normal index of 0–45). Individual chemistries are as follows: high density lipoprotein (HDL) = 28 versus normal expected values of 41–75 mg/dl; triglycerides = 165 versus normal expected values of 10–150 mg/dl; cholesterol = 132 versus normal expected values of 120–200 mg/dl; very low density lipoprotein = 33 versus normal expected values of 5–40 mg/dl; low density lipoprotein = 71 versus normal expected values of 80–130 mg/dl; cholesterol/HDL ratio = 4.71 versus normal expected values of 1.50–4.00.

All antiretroviral reference powders were supplied by the NIH AIDS Research and Reference Reagent Program. Methanol used to prepare stock solutions of ARVs was HPLC grade and was purchased from EMD Chemicals (Gibbstown, NJ, USA). HPLC grade water was also used, as needed for solubility, in the preparation of stock solutions and various vendors were utilized as a source (JT Baker, ACROS, EMD Chemicals).

Preparation of PT samples used to measure intra- & inter-assay variability

PT samples for intra-assay variability testing included duplicate samples sent to labs with separate identification numbers; intra-assay testing was conducted during two rounds of analysis. Each of these two rounds differed by which ARV was chosen for duplication. The inter-assay samples were prepared in bulk at "high" and "low" concentrations and used for testing during three consecutive rounds. High concentrations were targeted near the expected peak concentrations of the ARV and the "low" concentrations were five- to six-fold lower, with the exception of the NRTIs, which were only two fold lower due to the expected narrower concentration range. The lower concentrations were ten-fold or greater than concentrations at the lower limits of quantitation of these various laboratory methods. This design was intentional so as to maintain blinding of PT sample concentrations (e.g. neither trough concentrations nor peak concentrations were all within a single sample per level). Concentrations closer to expected trough values were found in other PT samples. One set of samples was shipped at the time of preparation and samples for subsequent testing were stored in bulk at −70 °C. These stored samples were sent to labs over the next 2 testing rounds (6 and 12 months post-preparation, respectively). Figure 1 displays the "value-added" schema, including the strategy used for measuring intra-assay variability and inter-assay variability.

Preparation of hemolytic & lipemic non-PT samples

A subsequent round included additional samples prepared in hemolyzed and lipemic plasma. These samples were spiked to achieve ARV concentrations that matched concurrent "control" PT samples, prepared in human plasma that was neither hemolyzed nor lipemic. Hemolyzed plasma was spiked at ARV concentrations matching those of two control samples; each matching sample contained different ARVs. Lipemic plasma was spiked at ARV concentrations matching those of three control PT samples; two of the three were for the same ARV but spiked at different concentrations. The control PT samples were scored; the hemolyzed and lipemic samples were included for investigational purposes only and testing was voluntary for each CPL. Both the PT (control) sample and the specificity (test) sample were analyzed simultaneously.

ARV long-term storage stability assessments

The next two rounds included prior PT samples as "stability test samples" for determining the long-term stability of selected ARVs in plasma stored at −70°C in one secure, monitored location for a designated period of time for each ARV. ARVs were selected based on the availability of initial results (control values) from CPLs and the currently expected CPL ARV testing selections. Stability test samples provided a testing window of 2.6–3.6 years. Stability testing was voluntary for each CPL.

PT sample prequalification, distribution & results report

Prior to each round, "pre-qualification" (to verify targeted concentrations) was completed with a goal of assuring concentrations were within 20% of the expected weighed-in value. If any prequalification mean had been found to be outside 20%, a root cause analysis and corrective action would have occurred to resolve the discrepancy. Notably, although specificity samples were not scored for accuracy, these were also prequalified to assure specificity at the pre-testing laboratory facility. Stability samples were also pre-tested but did not necessarily require prequalification due to the question of stability itself. The samples were then distributed to the participating CPLs for analysis. The CPLs reported results using the Laboratory Data Management System (version 5.6 – 6.1, Frontier Science Technology and Research Foundation); the results were compiled by the CPQA data management team. Additional information on the CPQA PT program operations has been reported in recent publications [2,17].

Analyses for method precision, specificity & stability during the analytical phase

While the PT program has offered testing for up to 21 ARVs, the analyses for intra- and inter-assay precision, method specificity and ARV stability focused on ARV results for atazanavir (ATV), darunavir (DRV), efavirenz (EFV), emtricitabine (FTC), lopinavir (LPV), nevirapine (NVP), raltegravir (RGV), ritonavir (RTV), and tenofovir (TFV) samples as included in the various testing measures and when three or more CPLs participated.

Calculation of intra- & inter-assay variability, effects of hemolysis & lipemia, & stability

Intra-assay variation was determined for each unique combination of CPL and ARV by calculating the relative percentage difference (RPD) between the 2 reported values: RPD = 100% × (largest value−smallest value)/(average of smallest and largest).

Inter-assay variation was determined for each CPL and ARV (at "high" and "low" concentrations) for samples included during 3 consecutive rounds, by calculating the relative standard deviation (RSD) of the 3 reported values: RSD = 100%*SD/mean, where SD is the standard deviation of the three values. Results for LPV were excluded due to a preparation error. Values from CPLs for specific ARVs were excluded if the CPL did not test and report for all three rounds.

The effect of hemolysis and lipemia on the accuracy of the CPL or ARV assay was determined by calculating the percentage difference (PD) between the results for the test and control samples: PD = 100% X (test value−control value)/(control value). These

supplemental results were summarized as the number of values where PD <5%, 5–10%, and >10% with no specific acceptability target.

The percentage recovery (PR) for stability determinations was determined by dividing the median recovery value for the analyte stability sample value (test) by the median of the initial testing of the PT samples as the time zero or control value, where stability years is equal to the time between testing of test versus control. All initial median control values were within 10% of their nominal concentrations. One hundred percentage recovery would indicate that ideal stability was achieved, whereas values below 100% would indicate degradation of the analyte in the sample. ARV-specific results were excluded when the CPL did not pass proficiency for the specific ARV.

Statistical evaluations

When summarizing the intra- and inter-assay variability across CPLs for an ARV or across ARVs for a CPL, the median value was chosen due to the small numbers of participating CPLs per ARV as well as the smaller number of some ARV results. To evaluate associations between intra- or inter-assay variability and CPL, ARV or concentration level, the Kruskal-Wallis (KW) test was performed independently for each variable at the 5% level of significance. For each variable exhibiting statistically significant differences among RPD or RSD via KW: [1] outcome measures (RPD or RSD) were ranked, [2] a one-way analysis of variance model (ANOVA) was fit to the ranks, and [3] all pairwise comparisons were conducted using the Tukey test with 5% family-wise error rate. For lipemia and hemolysis effects, each CPL and ARV was considered independently, due to potential assay differences affecting specificity (e.g. each CPL validated and applies a unique method). For ARV stability testing, results across laboratories were pooled and the median value was chosen as the final recovery value as the number of values per ARV were limited and their distribution could therefore not be assessed. Stability test values (or control values) were excluded from analysis when a laboratory failed proficiency for the analyte of interest. The KW test was used to compare the medians of the original or control values to the stability testing values. Analyses were conducted using Minitab 16 Statistical Software (Minitab, Inc., version 16.1.0).

Results & discussion

Prequalification testing

Prequalification testing provided support for the accurate preparation of PT samples with one exception. The bulk preparations for the blinded inter-assay precision monitoring did not include LPV as an analyte. Route cause analysis uncovered a failure to include the analyte in the final preparation schema and the cause of this failure was found to be a spreadsheet omission during spreadsheet reorganization. Since all aliquots of PT were already distributed to vessels for the first of three rounds of testing, a decision to drop LPV from the inter-assay precision monitoring was made. Remediation included other steps taken to include LPV in the three testing rounds as PT samples alone (not precision).

Intra-assay variability

Intra-assay variability is summarized in Table 1. Intra-assay variability, as measured by RPD between duplicates, ranged from 0.28 to 12.5% with a median RPD of 1.61%. Tenofovir demonstrated the highest median RPD with a value of 3.38%. CPL 4 had the highest median RPD with a value of 4.02. Medians values, whether summarized by ARV or by CPL, were 4%. Neither CPL nor ARV had a significant effect on RPD (KW p=0.204 and 0.529, respectively). Of the 49 calculated RPD, 46 (94%) were 10%. The design of intra-assay precision monitoring was limited by a PT sample number limit of 5, the need to vary ARV concentrations of PT samples to test a relevant therapeutic concentration range during each round, and the PT requirement to maintain blinding. Thus only two samples were used to test intra-assay variability and intra-assay variability is measured by RPD rather than RSD. However, a few design factors provide unique strengths: samples were blinded, not all samples contained duplicate concentrations of all contained ARV, and in some samples a similar (but not duplicate) concentration was included. For example, while sample 1 and sample 2 contain duplicate concentrations of ATV, EFV, RGV and TFV, concentrations of DRV, FTC, and RTV are not identical in these two samples. These results provide confidence of acceptable and continued precision or reproducibility in the CPL sample preparation and analysis within an applied, previously validated assay.

Inter-assay variability

Inter-assay variability, as measured by RSD of three replicates at high and low concentrations across three rounds is shown in Table 2. The range of all RSD was 0.38–23.3%. The median RSD for all the data collected is 3.69%. The range of RSD for CPL medians is 2.78–13.5% and the median range for ARV RSD (high or low) is 1.84–5.03%. Neither level (high vs. low concentration) nor ARV had a significant effect on RSD (p=0.251 and p=0.521, respectively). However, the individual CPL did have a significant effect on RSD (p=0.012), and by rank comparison and Tukey test, one CPL was different (greater RSD) from another CPL (smallest RSD). Ninety eight percentage of inter-assay results obtained in three consecutive rounds revealed less than 10% variability and only two CPL showed a value greater than 15% (both for FTC). Both of these higher values are attributable to a proficiency failure of one or more samples (not within 20% of the target) but not always failure for the analyte overall. As mentioned in the background section, the effect of ARV concentrations on inter-laboratory precision measurements have been studied and reported [3–7]. While some investigators have identified a tendency for lower concentrations to show more variability, others have reported no statistical differences. Since the "low" concentration PT samples were intentionally not targeted to the expected trough concentration so as to maintain a blinded design, the insignificance of the variability of high versus low concentrations cannot be compared within the same context. But overall, these blinded measures of inter-assay variation or precision provide support for the continued inter-assay reproducibility during the analysis of trial specimens conducted over 1 year of time at CPLs.

Hemolysis & lipemia

Calculated PDs between control and specificity samples are shown in Tables 3 & 4. The control sample was scored for proficiency ($< \pm 20\%$ of final target) whereas the test sample was not. Hemolyzed sample results are summarized in Table 3. For results of hemolyzed samples, most values (87%) exhibited PD from the lab control value below 10% (in absolute value). EFV, FTC, LPV and NVP were the ARVs where differences were largest. Of 105 results for lipemic samples, 15 (14%) differed more than 10% from control samples. Results are summarized in Table 4. Please note that with the exception of NVP, two specificity concentration values were available for each ARV and the CPL number was noted as outside the 10% PD window only when both values showed this difference. Reported lipemic results for DRV, RGV, TFV, and NVP from a single CPL accounted for seven of the 15 results (not shown). Some of the chromatographic methods employed by CPLs for ARV testing employ detection methods less specific than that of tandem mass spectrometry such as single quadrupole detection or ultraviolet detection. In addition, the expense and availability of some isotopic internal standards can be prohibitive for those CPLs using tandem mass spectrometry due to budgeting limitations. Therefore, testing for the effect of common specimen abnormalities on these chromatographic assays is important. Patients receiving HIV-1 protease inhibitors, or who have other concomitant diseases such as diabetes, may develop higher levels of triglycerides causing study samples obtained to be lipemic. Hemolysis can occur during sample collection, handling and processing or in patients co-infected with hepatitis C virus receiving ribavirin. These and other plasma abnormalities commonly found in HIV infected and HIV co-infected patients may cause matrix effects in the chromatographic sample preparation or detection processes. In clinical laboratory method validation, these types of abnormalities are tested [18] and more recently, as previously mentioned, this is been a focus for method validation requirements in the EMEA [8]. Many of the bioanalytical methods employed by the CPLs have been validated using the FDA 2001 guidance prior to the wide-spread discussion of these two particular specimen anomalies. Therefore, it was decided that results for these included specimens would not be scored at first introduction, but rather serve as a tool for method remediation as needed with no specific target range for acceptance. Overall, hemolysis and lipemia had little effect on the accuracy of the reported ARV values with the exception of lipemic samples tested by one CPL for multiple ARVs. For this CPL, and the CPLs where some ARVs may have been consistently affected, the inclusion of abnormal samples was helpful in defining specificity issues of the employed bioanalytical method. CPQA was able to alert these CPLs of potential specificity issues and provide recommendations for investigation and/or remediation. Future examples of possible specificity testing may include plasma with total protein content less than the normal range, drug metabolite presence, icterus which can occur in patients receiving ATV, or elevated hepatic enzymes such as those that increase in co-infection with hepatitis C. It should be noted, however, that it is important to include a control sample when testing for specificity failures. Placing control samples within the normal PT testing samples is efficient and effective, since blinding can be maintained by varying placement. This approach is limited since the specimens were not derived from patients receiving ARVs and therefore do not take into account dilution or addition of ARV from hemolysis of red cells as well as patient-specific lipemic conditions.

Stability

Ten laboratories participated in stability testing for select ARVs and results were determined for five ARVs. Of the 98 stability results displayed in Table 5, five results were excluded due to inaccuracy as defined by PT failure (<80% ARV sample concentrations within ±20%) or identified as outliers by the Grubbs test. The recovery results (using medians) for all five ARV exhibit 97% recovery for 70% of the stability samples tested and range overall from 88% to 102% recovery. DRV was shown to be stable for 2.58 years in plasma stored at −70±7°C. FTC, NFV, RGV and TFV were shown to be stable under the same conditions for 3.50–3.58 years. The KW test results comparing individual original results to stability test results indicated there were no significant differences over time. Across all ARVs, 65% of the median ARV stability results exhibited 97% recovery. This finding is not unexpected, as multiple CPL measures should minimize overall bias, and therefore, provide a better test of stability results. Since the laboratories testing the stability samples also tested the round PT sample for the same analytes, stability results from CPL failing proficiency were excluded and therefore stability results were qualified as accurate by PT. The stability testing was not intended to replace the regulatory guidance recommendations for stability testing, but to supplement existing laboratory data and allow support for the analysis of clinical samples after long-term storage at −70°C. Four limitations in this stability testing approach are noted. First, successful stability is not derived against freshly prepared samples. This can be altered in the future stability testing procedures using fresh PT samples as controls, however "fresh" would require an assumption of up to three months of stability given the processes of the PT program itself (e.g., preparation, prequalification, shipments and allowance of time for CPL testing). Laboratories would still need to provide analyte stability in-house according to its applicable regulatory guidance. Second, the availability of ARV in the CPQA PT program or validated methods at CPLs to test newer ARV drugs is limited. ARVs that have not yet been added by the CPQA PT program include elvitegravir, rilpivirine, and cobicistat; ARVs that did not yet meet criteria for this stability testing approach include etravirine and maraviroc." Stability testing of these agents will be considered in future panel designs. Third, although ARV stability was tested in human plasma, the specimens were not derived from patients receiving ARVs and therefore do not take into account conversion of metabolites as well as patient-specific and trial-specific stability. Lastly, other long-term environmental storage conditions ought to be considered for resource-limited settings, such as −20°C storage.

Conclusion

By including investigational samples in proficiency testing rounds, these samples were analyzed with routine or control PT samples, and provided valuable laboratory performance information. The CPQA PT program assured that intra-assay and inter-assay precision or variability continue meet 15% acceptability criteria after method validation (albeit with fewer sample number), when performing assays for clinical protocols. In addition, the specificity of these methods when lipemia or hemolysis was present was also confirmed for most methods and allowed CPQA to communicate with specific CPLs as to potential specificity issues. Lastly, while stability is most often defined through internal mechanisms, the CPQA stability testing approach offered a shared, robust system of stability testing to

multiple laboratories quantifying ARVs in plasma specimens stored frozen for 2.5–3.6 years.

Future perspective

"Added-value" PT designs will continue to be offered by the CPQA program, as an efficient, economic approach to assure high quality data is obtained within NIAID HIV pharmacology research. As new drugs, including drugs given to treat tuberculosis and hepatitis or common comorbidities, are added to the NIAID network agenda and protocols, the need to follow FDA and in-country method validation guidance will remain essential. For the CPQA PT program it will be critical to continue to test and monitor the accuracy and inter-laboratory variability of drug result measurements, the continued CPL precision (intra- and inter-assay) of the bioanalytical methods used, analyte stability in specimens, and the effect of known or newly discovered endogenous substances on the results. A future improvement or proof of concept for the stability testing approach could include freshly prepared controls. Some specificity testing goals could also become part of the scored proficiency testing program. As participating labs have demonstrated proficiency and the CPQA "added-value" PT designs have proven to be of meaningful, this approach will optimize the value future proficiency testing.

Acknowledgments

Key terms

Proficiency testing	Test of accuracy for specific analytes measured by a laboratory; the Clinical Laboratory Improvement Act (CLIA) requires laboratories to show proficiency of their testing methods for purposes of maintaining licensure
The Clinical Pharmacology Quality Assurance Program	Maintains a multifaceted approach to supporting the quality of HIV clinical research conducted by various trial networks. Two laboratory aspects are bioanalytical method validation and proficiency testing
Clinical Pharmacology Laboratories	Provide bioanalytical support for multiple Clinical Trial Networks such as the AIDS Clinical Trial network (ACTG), the International Pediatric, Maternal, and Adolescent AIDS Clinical Trial network (IMPAACT), the HIV Prevention Trial Network (HPTN) and the Microbicides Trial Network (MTN)
"Value added"	Can be incorporated by purposely designing PT samples to also assess measures other than accuracy or by including non-PT samples for voluntary testing

| **Blinding** | Term used in proficiency testing that indicates that the true or estimated value of the measured constituent is not known to the laboratory |

References

Papers of special note have been highlighted as:

• of interest; •• of considerable interest

1. US DHHS. Medicare, Medicaid and CLIA programs; regulations implementing the Clinical Laboratory Improvement Amendments of 1988 (CLIA)–HCFA. Final rule with comment period. Fed Regist. 1992; 57(40):7002–7186. [PubMed: 10170937]

2. DiFrancesco R, Rosenkranz SL, Taylor CR, et al. Clinical Pharmacology Quality Assurance (CPQA) Program: Longitudinal Analysis of Antiretroviral (ARV) Proficiency Testing for International Laboratories. Ther Drug Monit. 2013; 35(5):631–642. [PubMed: 24052065]

3. Holland DT, DiFrancesco R, Stone J, Hamzeh F, Connor JD, Morse GD. Adult and Pediatric AIDS Clinical Trials Group Pharmacology Laboratory Committees, Pediatric AIDS Clinical Trials Group. Quality assurance program for clinical measurement of antiretrovirals: AIDS clinical trials group proficiency testing program for pediatric and adult pharmacology laboratories. Antimicrob Agents Chemother. 2004; 48(3):824–831. [PubMed: 14982771]

4. Holland DT, DiFrancesco R, Connor JD, Morse GD. Quality assurance program for pharmacokinetic assay of antiretrovirals: ACTG proficiency testing for pediatric and adult pharmacology support laboratories, 2003 to 2004: a requirement for therapeutic drug monitoring. Ther Drug Monit. 2006; 28(3):367–374. [PubMed: 16778721]

5. Aarnoutse RE, Verweij-van Wissen CP, van Ewijk-Beneken Kolmer EW, et al. International interlaboratory quality control program for measurement of antiretroviral drugs in plasma. Antimicrob Agents Chemother. 2002; 46(3):884–886. [PubMed: 11850280]

6. Droste JA, Aarnoutse RE, Koopmans PP, Hekster YA, Burger DM. Evaluation of antiretroviral drug measurements by an interlaboratory quality control program. J Acquir Immune Defic Syndr. 2003; 32(3):287–291. [PubMed: 12626888]

7. Burger D, Teulen M, Eerland J, Harteveld A, Aarnoutse R, Touw D. The International Interlaboratory Quality Control Program for Measurement of Antiretroviral Drugs in Plasma: a global proficiency testing program. Ther Drug Monit. 2011; 33(2):239–243. [PubMed: 21383652]

8•. EMA. EMA. Committee for Medicinal Products for Human Use; London, UK: 2011. Guideline on Bioanalytical Method Validation. Specificity in the presence of common abnormalities is required by the EMA guidelines. Hemolysis is one plasma abnormality that these guidelines specifically list

9•. US FDA. Guidance for Industry: Bioanalytical Method Validation. US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research; Rockville, MD, USA: 2001. The US FDA requires that the intra-assay and inter-assay precision of a validated bioanalytical method be ≤15% and that stability be proven for the intended storage conditions

10. Savoie N, Garofolo F, van Amsterdam P, et al. White Paper on Recent Issues in Regulated Bioanalysis & Global Harmonization of Bioanalytical Guidance. Bioanalysis. 2010; 2(12):1945–1960. [PubMed: 21110739]

11. Berube E, Taillon M, Furtado M, Garofolo F. Impact of sample hemolysis on drug stability in regulated bioanalysis. Bioanalysis. 2011; 3(18):2097–2105. [PubMed: 21942520]

12. Dimeski G. Interference Testing. Clin Biochem Rev. 2008; 29(1):S43–S48. [PubMed: 18852856]

13. Lindegardh N, Hanpithakpong W, Kamanikom B, et al. Major pitfalls in the measurement of artemisinin derivatives in plasma in clinical studies. J Chromatogr B Analyt Technol Biomed Life Sci. 2008; 876(1):54–60.

14. Chin C, Zhang ZP, Karnes HT. A study of matrix effects on an LC/MS/MS assay for olanzapine and desmethyl olanzapine. J Pharm Biomed Anal. 2004; 35(5):1149–1167. [PubMed: 15336361]

15. Hughes N, Bajaj N, Fan J, Wong E. Assessing matrix effects of hemolyzed samples in bioanalysis. Bioanalysis. 2009; 1(6):1057–1066. [PubMed: 21083074]

16. Cembrowski GS, Carey RN. Adding value to proficiency testing programs. Clin Chem. 2000; 46(1):7–8. [PubMed: 10620565]

17•. DiFrancesco R, Tooley K, Rosenkranz SL, et al. Clinical Pharmacology Quality Assurance for HIV and Related Infectious Diseases Research. Clin Pharmacol Ther. 2013; 93(6):479–482. The Clinical Pharmacology Quality Assurance and Quality Control program provides support to ensure the integrity of clinical trial pharmacology results. A proficiency testing program specifically ensures that participating laboratories are accurately quantitating antiretrovirals of current interest. [PubMed: 23588323]

18. Kazmierczak, SC. Hemolysis, Lipemia, and High Bilirubin: Effect on Laboratory Tests. In: Dasgupta, A.; Sepulveda, JL., editors. Accurate Results in the Clinical Laboratory: A Guide to Error Detection and Correction. Elsevier Inc; Waltham MA, USA: 2013. p. 53-62.

Executive Summary

Background

- Proficiency testing (PT) for ARVs in therapeutic drug monitoring (TDM) and selected clinical trial protocols provides assurance of laboratory accuracy and measures variability between participating laboratories.

Intra- & inter-assay precision

- The use of PT to assess and verify intra- and inter-assay precision in a blinded manner can provide assurance that the bioanalytical methods continue to meet quality standards while clinical research trial samples are analyzed by previously validated methods.

Specificity

- Investigational PT samples can be provided to assure the absence of deleterious effects by endogenous substances and offer valuable information for quality assurance or needs for assay remediation in laboratories engaged in clinical HIV antiretroviral research.

Stability

- PT programs can be used to measure the stability of its testing analytes stored in specific environmental conditions by retesting prior PT samples as investigational or "stability" samples; testing across multiple laboratories may eliminate bias and provide a unique opportunity for efficient collaborative work.
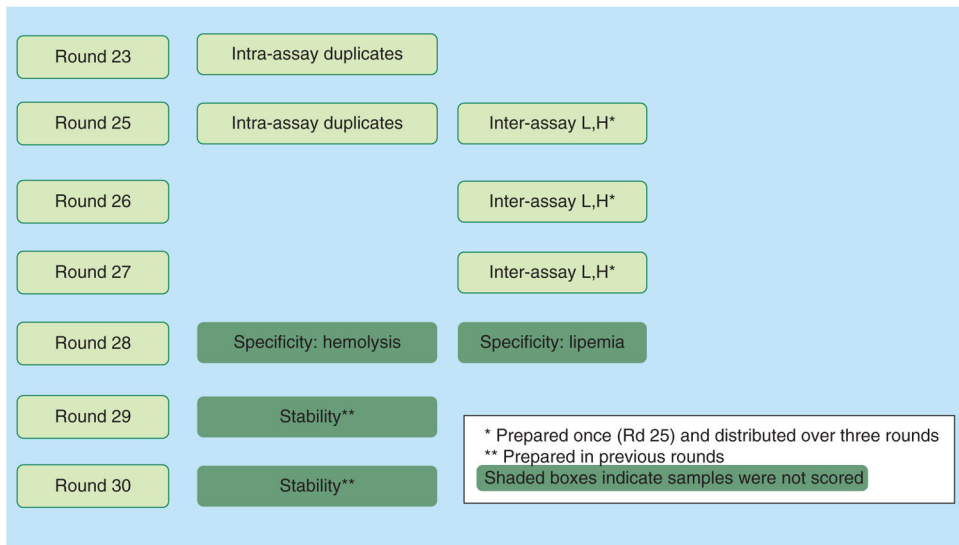
Figure 1.
Adding more value to proficiency testing rounds.

Table 1

Intra-assay variation measured by duplicated concentrations in two proficiency samples tested within a round. Relative percentage difference = 100% × ([largest value–smallest value]/average of smallest and largest).

ARV←CPL↓	ATV	DRV	EFV	FTC	LPV	RGV	TFV	Median CPL RPD	Min/Max
Target concentrations (ng/ml)	*1200*	*3500*	*1500[†]*	*600/1250[†]*	*1200*	*800/900[†]*	*120[†]*		
1	3.15	4.54	3.57	0.90	0.86	-	7.09	**3.36**	0.86/7.09
2	1.22	0.70	1.65	0.64	0.56	0.28	11.50	**0.70**	0.28/11.50
3		-	2.51	-	-	0.45	-	**1.48**	0.45/2.51
4	12.5	-	3.51	1.61	8.69	0.75	4.53	**4.02**	0.75/12.5
5	1.06	0.76	3.57	-	5.14	-	-	**2.32**	0.76/5.14
6	1.40	0.39	3.49	5.76	6.87	-	0.97	**2.45**	0.39/6.87
7		1.17	4.92	1.32	-	-	0.76	**1.25**	0.76/4.92
8	1.29	-	0.69	-	0.50	-	-	**0.69**	0.50/1.29
9		4.44	1.48	3.62	-	0.57	-	**2.55**	0.57/4.44
10	0.76	10.10	1.26	6.97	2.32	10.80	2.23	**2.32**	0.76/10.80
Median ARV RPD	**1.29**	**1.17**	**3.00**	**1.61**	**2.32**	**0.57**	**3.38**	Overall Median	1.61
Min/Max	**0.76/12.5**	**0.39/10.10**	**0.69/4.92**	**0.64/6.97**	**0.50/8.69**	**0.28/10.80**	**0.76/11.50**		

[†] These duplicate variability was tested in both rounds for these drugs.

The laboratory provided results for both rounds of duplicate testing and the first RPD is reported.

ARV: Antiretroviral; ATV: Atazanavir; CPL: Clinical pharmacology laboratory; DRV: Darunavir; EFV: Efavirenz; FTC: Entricitabine; LPV: Lopinavir; RGV: Raltegravir; RPD: Relative percentage difference; TFV: Tenofovir.

Table 2

Inter-assay variability (RSD) measured across three proficiency testing rounds and summarized by clinical pharmacology laboratory, antiretroviral and concentration level (n = 1 result per level per analyte per laboratory per round).

CPL	ATV		DRV		EFV		FTC		RGV		RTV		TFV		Median CPL RSD		CPL Min/Max
	H	L	H	L	H	L	H	L	H	L	H	L	H	L	H	L	H&L
Target concentration (ng/ml)	7000	1500	7000	1500	9000	1500	2500	1250	4500	900	2500	500	240	120			
1	3.29	2.79	1.54	3.32	4.77	1.02	2.84	2.77	-	-	4.34	3.70	2.25	2.46	3.07	2.78	1.02/4.77
2	2.44	3.50	3.93	3.67	3.67	1.84	3.39	1.98	7.04	4.06	2.12	4.53	5.89	3.04	3.67	3.50	1.98/7.04
3	-	-	-	-	6.76	7.07	-	-	6.10	1.95	2.81	4.63	-	-	6.10	4.63	1.95/7.07
4	-	-	-	-	-	-	2.79	21.5[†]	-	-	-	-	5.08	5.59	3.94	13.5	2.79/21.5
5	5.05	1.17	4.32	2.62	6.07	3.52	-	-	-	-	6.48	7.48	-	-	5.56	3.07	1.17/7.48
6	3.05	3.45	4.20	2.96	1.57	2.78	**23.3**[‡]	8.79	-	-	3.22	5.32	-	-	3.22	3.45	1.57/**23.3**
7	3.09	6.24	5.99	8.46	6.85	5.03	7.01	7.28	-	-	4.13	5.20	4.98	4.71	5.49	5.72	3.09/8.46
8	4.08	3.67	-	-	0.38	0.96	-	-	-	-	3.10	3.21	-	-	3.10	3.21	0.38/4.08
9	4.95	3.13	5.39	3.09	5.24	0.78	5.38	3.09	2.64	3.67	4.82	1.83	2.16	3.36	4.95	3.09	0.78/5.39
10	7.45	7.25	7.85	10.6	2.35	0.65	4.12	3.81	2.75	2.37	-	-	7.13	6.45	5.62	5.13	0.65/10.6
Median ARV, RSD	3.68	3.47	4.32	3.32	4.77	1.84	4.12	3.81	4.42	3.02	3.67	4.58	5.03	4.03	Median H	Median L	Overall Median
Min/Max	1.17/7.25		1.54/10.6		0.38/7.07		2.77/**23.3**		2.75/7.04		1.83/7.48		2.16/7.13		4.32	3.48	3.69

[†] Laboratory failed 1/3 rounds for proficiency; please note that normal QC will have passed.

[‡] Laboratory was outside 20% acceptance for 2/3 values but still passed proficiency at 80%; please note that normal QC will have passed.

Values >15% are formatted as bold and underlined.

ARV: Antiretroviral; ATV: Atazanavir; CPL: Clinical pharmacology laboratory; DRV: Darunavir; EFV: Efaviren; FTC: Emtricitabine; H: high; L: low; LPV:Lopinavir; RGV: Raltegravir; RPD: Relative percentage difference; RSD: Relative standard deviation; RTV: Ritonavir; TFV: Tenofovir.

Table 3

Specificity: percentage difference ranges in antiretroviral analytes measured when hemolysis is present %PD = ([hemolyzed value − control value]/control value) ×100.

Analyte	N (values reported)	PD <±5 n(%)	PD ± 5–10 n(%)	PD >±10 n(%)	CPL
ATV	5	5 (100%)	0 (-)	0 (-)	-
DRV	7	4 (57%)	3 (43%)	0 (-)	-
EFV	9	6 (67%)	2 (22%)	1 (11%)	8
FTC	4	2 (50%)	1 (25%)	1 (25%)	9
LPV	8	5 (63%)	2 (25%)	1 (12%)	6
NVP	5	1 (20%)	1 (20%)	3 (60%)	2, 6, 7
RGV	5	3 (60%)	2 (40%)	0 (-)	-
RTV	8	5 (63%)	3 (37%)	0 (-)	-
TFV	5	4 (80%)	0 (0%)	1 (20%)	7
Total	56	35 (62%)	14 (25%)	7 (13%)	-

%PD = ([hemolyzed value − control value]/control value) ×100.

N = Total number of reported results; n = Number in subset of results within the PD range.

ARV: Antiretroviral; ATV: Atazanavir; CPL: Clinical pharmacology laboratory; DRV: Darunavir; EFV: Efaviren; FTC: Emtricitabine; LPV: Lopinavir; NVP: Nevirapine; RGV: Raltegravir; RPD: Relative percentage difference; RTV: Ritonavir; TFV: Tenofovir.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 4

Specificity: percentage difference in antiretroviral measured when lipemia present.

Analyte	N (values reported)	PD <±5 n(%)	PD ± 5–10 n(%)	PD >±10 n(%)	CPL[†]
ATV	11	5 (45%)	5 (45%)	1 (10%)	-
DRV	14	8 (57%)	4 (29%)	2 (14%)	7
EFV	16	12 (75%)	3 (19%)	1 (6%)	-
FTC (1.6)	8	6 (75%)	2 (25%)	0 (-)	-
LPV	16	11 (69%)	4 (25%)	1 (6%)	-
NVP	5	2 (40%)	1 (20%)	2 (40%)	2,7[‡]
RGV	10	7 (70%)	0 (-)	3 (30%)	7
RTV	15	7 (47%)	5 (33%)	3 (20%)	9
TFV	10	8 (80%)	0 (-)	2 (20%)	7
Total	105	66 (63%)	24 (23%)	15 (14%)	-

[†]Only if CPL consistently >±10 for two samples.

[‡]By design NVP not tested in duplicate for lipemia, thus only 1 value >±10.

%PD = (|lipemic value − control value|/control value) ×100

ARV: Antiretroviral; ATV: Atazanavir; CPL: Clinical pharmacology laboratory; DRV: Darunavir; EFV: Efaviren; FTC: Emtricitabine; LPV: Lopinavir; N: Number; NVP: Nevirapine; RGV: Raltegravir; RTV: Ritonavir; TFV: Tenofovir.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 5

Stability results.

ARV	N CPL testing original sample	N[†] excluded CPL testing original sample	Median for original results	N CPL testing stability sample	N[†] excluded CPL testing stability sample	Median for stability results	Median, % recovery	Storage duration	Kruskal–Wallis test, p-value
DRV	6	1	1547	8	1	1581	102%	2.58	0.606
DRV	7	0	3955	9	0	3920	99%	2.50	0.634
FTC	3	0	611	4	1	606	99%	3.58	1.00
FTC	5	0	587	5	1	576	98%	3.50	0.753
FTC	4	1	823	4	1	804	98%	2.50	0.564
FTC	7	0	184	5	1	179	97%	2.50	0.464
NVP	6	0	757	7	0	706	93%	3.50	0.568
NVP	6	0	1636	7	0	1589	97%	3.50	0.391
NVP	7	0	3140	5	0	2956	94%	2.50	0.062
RGV	4	0	346	6	0	350	101%	3.50	0.831
RGV	4	0	818	6	0	807	98%	3.50	0.670
RGV	3	0	1745	6	0	1539	88%	2.58	0.121
RGV	4	0	1004	6	0	989	98%	2.50	0.831
TFV	3	0	221	4	0	203	92%	3.58	1.00
TFV	4	0	126	6	0	113	90%	3.50	0.286
TFV	5	0	148	4	0	146	99%	2.58	0.327
TFV	6	0	129	6	0	120	93%	2.50	0.229

[†] Values excluded when Grubbs test identified as outlier or laboratory failed proficiency for the ARV Thinner line segregates ARV stability testing occurring in more than one round.

% recovery = 100% × (stability test value/final proficiency test target value in original testing round).

ARV: Antiretroviral; ATV: Atazanavir; CPL: Clinical pharmacology laboratory; DRV: Darunavir; EFV: Efaviren; FTC: Emtricitabine; LPV:Lopinavir; N: Number; NVP: Nevirapine; RGV: Raltegravir; RPD: Relative percentage difference; RTV: Ritonavir; TFV: Tenofovir.