

**ASVCP quality assurance guidelines: External quality assessment and comparative testing for reference and in-clinic laboratories**

**Version 1.0 (7/2015)**

Developed by the American Society for Veterinary Clinical Pathology (ASVCP) Quality Assurance and Laboratory Standards (QALS) Committee

Melinda S. Camus<sup>1</sup>, Bente Flatland<sup>2</sup>, Kathleen P. Freeman<sup>3</sup>, Janice A. Cruz Cardona<sup>4</sup>

<sup>1</sup>Department of Pathology, College of Veterinary Medicine, University of Georgia, Athens, GA, USA; <sup>2</sup> Department of Biomedical and Diagnostic Sciences, College of Veterinary Medicine, University of Tennessee, Knoxville, TN, USA; <sup>3</sup>IDEXX Laboratories Ltd., Wetherby, West Yorkshire, UK; <sup>4</sup> IDEXX Laboratories Inc., Houston, TX, USA

**Contents**

1. Purpose and Scope.....	2
2. Introduction .....	2
3. Definitions .....	3
4. External quality assessment.....	6
4.1 Overview .....	6
4.2 Benefits of EQA participation.....	7
4.3 Limitations of EQA:.....	8
4.4 Recommendations for participating in EQA programs.....	8
4.5 Handling and processing testing materials.....	9
4.6 Evaluation of EQA performance.....	10
4.7 Managing unacceptable results .....	10
5. Other Comparative Testing.....	12
5.1 Other comparative testing -- Overview .....	12
5.2 Frequency and design of comparative testing.....	12
5.3 Design of comparative testing events.....	13
5.4 Specimens used for comparative testing.....	14
5.5 Basis for determining results comparability .....	15
5.6 Clinical significance of result comparability.....	16
5.7 Reasons for non-comparable results .....	17
6. Acknowledgements.....	17

7. References .....	17
Appendix A.....	20
Appendix B.....	21
Appendix C.....	24
Appendix D.....	28
Appendix E .....	30

## 1. Purpose and Scope

Laboratory quality management programs should include different types of comparative testing. Comparative testing may include participation in formal external quality assessment (EQA) programs (also known as proficiency testing, PT) and less formal comparisons between testing sites (e.g., veterinary clinics and reference laboratories). The purpose of formal comparative testing programs (EQA/PT) is for a laboratory to monitor its performance relative to a peer group using the same methods or relative to known values of reference materials. Less formal comparative testing may help “check” a testing site’s (e.g., veterinary clinic’s) performance against another laboratory or field method. The emphasis of comparative testing is assessment of bias.

The purpose of this document is to educate providers of veterinary laboratory diagnostic testing in any setting about comparative testing. These guidelines will define, explain, and illustrate the importance of a multi-faceted laboratory quality management program which includes comparative testing. The guidelines will provide suggestions for implementation of such testing, including which samples should be tested, frequency of testing, and recommendations for results interpretation. Examples and a list of vendors and manufacturers supplying control materials and services to veterinary laboratories are also included.

These guidelines are designed to complement other ASVCP quality guidelines, which can be found on the ASVCP website (<http://www.asvcp.org/pubs/qas/index.cfm>).

## 2. Introduction

Quality assurance (QA) is the series of procedures employed to evaluate all laboratory systems and processes in order to identify problems, correct them, and continue to enhance the level of performance in all phases of testing. QA includes quality planning, implementation, monitoring, and assessment.<sup>1,2</sup> This is the portion of quality management that seeks to provide confidence to laboratory staff and clients that quality requirements for analyzed samples are being met.<sup>3</sup> QA incorporates monitoring and detection of problems arising during routine operation of a laboratory and includes quality control (QC or “running controls”) as well as other procedures both within and outside of the laboratory.

Comparative testing is a QA procedure that can be external or internal to the laboratory and typically involves comparison of results between two or more instruments or methods for purposes of analytical performance assessment. There are different types of comparative evaluations, including formalized EQA and PT programs to which laboratories can subscribe.

The term PT is sometimes used specifically to refer to programs that compare measured results to target values established by reference methods, often for compliance with governmental regulations. In the United States, the terms EQA and PT may be used interchangeably.<sup>4,5</sup> Comparative testing for QA purposes can also encompass independent assessments carried out by individual laboratories. Comparative testing can be done across different instruments, different methodologies, and across laboratory systems, depending on the intended goal of the evaluation. All of these are considered separate from routine analysis of QC samples and do not replace the need to perform daily in-laboratory QC.

### 3. Definitions

**Acceptance criteria** – In the context of comparative testing, these are performance characteristics defined by an individual laboratory and/or in expert guidelines indicating that 2 methods/instruments being compared are producing equivalent results based on the intended use of the test result. Acceptance criteria may be based on analytical quality requirements (quality specifications) and can be derived from allowable total error ( $TE_a$ ), biologic variation data, standard deviation index (SDI), number of standard deviations from a peer group mean, or other performance parameters.

**Analyte** - The substance or chemical undergoing analysis or measurement. It can be used synonymously with *variable*, but is distinct from *measurand*.<sup>6</sup>

**Bias (inaccuracy)** – Total systematic error, which includes constant and proportional bias. Bias is the difference between the measured result and some measure of the “true” value (e.g., as measured by a reference method or as defined by a known standard). The term *bias* has a specific meaning in the statistical t-test and in difference plot analysis, where bias (expressed in analyte units) equals the difference between the mean values of two methods being compared or the average of all the differences between the paired sample values. Bias may also be expressed as a percentage according to the formula

$$Bias\% = \frac{Mean_{measured} - Mean_{target}}{Mean_{target}} \times 100$$

In clinical pathology laboratories, best practice dictates that target means be based on data from method comparison to a true reference method (“definitive” method) or known concentration of certified reference material. Target means may also be based on peer group means.<sup>1</sup>

**Bias, constant** – When the degree of systematic error remains the same over the range of analyte concentrations (i.e., results of one method are consistently above or below another method).<sup>1</sup>

**Bias, proportional** – When the magnitude of systematic error changes as the analyte concentration changes. Often, error increases as the analyte concentration increases, but the reverse may also be true.<sup>1</sup>

**Biological variation** - Fluctuation of analyte concentration around a homeostatic set point or according to daily, monthly, or seasonal rhythms.<sup>7</sup>

**Coefficient of Variation (CV)** - A measurement of imprecision (random error); mathematically, CV is standard deviation expressed as a percentage of the mean.<sup>3</sup>

**Commutability** - Equivalence of the mathematical relationships between the results of different measurement procedures for a reference material and for representative samples from healthy and diseased individuals. In practical terms, the property of commutability refers to the fact that a material interacts with the test system in a manner similar to patient samples.<sup>8</sup>

**Comparability** - Agreement between patient results obtained for a measurand using different measurement procedures within a health care system or different instruments within or outside the laboratory. Results are considered comparable if differences do not exceed an established critical value based on defined acceptance criteria.<sup>9</sup>

**Comparative Testing** - A QA procedure in which measurement results from two or more instruments or methods are compared to each other for purposes of analytical performance assessment. Comparative testing can be a component of formal EQA/PT programs or can be carried out independently and internally within a laboratory or network of laboratories.

**External Quality Assessment (EQA)** - Interlaboratory comparisons and other performance evaluations that may extend throughout all phases of the testing cycle, including interpretation of results. Peer comparison and comparison with known values of reference materials are types of EQA.<sup>4</sup>

**EQA/PT sample, testing item, test material, or check sample panel** - A sample containing measurands of undisclosed concentrations or compositions sent to a participating laboratory in order to assess the laboratory's testing competency.<sup>4</sup>

**Matrix Effect** - Refers to the influence of physical and/or chemical properties of a given sample or control material, other than the measurand, on the quantitation of that measurand and thus on the reported value according to a particular methodology. Causes of matrix effect include turbidity, viscosity, protein composition, and pH, among others. Matrix effects are a potential cause of non-commutability of samples across different test systems and must be distinguished from lack of results comparability due to bias.<sup>4</sup>

**Matrix Free Material** - Testing material (reagent, control, sample, etc.) that has minimal, if any, matrix effect.

**Measurand** - A particular quantity subject to measurement under specified conditions (e.g. the enzymatic activity of alkaline phosphatase at 37°C).<sup>4</sup>

**Metrological traceability** - Property of a measurement result whereby the result can be related to a reference material or method through a documented unbroken chain of calibrations, each contributing to the measurement of uncertainty.<sup>10</sup>

**Parameter** - A measurable factor that influences the changes of an analyte or variable (e.g. patient age in alkaline phosphatase activity; incubation time in serology assays).<sup>6</sup>

**Peer comparison** - A type of external quality assessment where laboratory results are compared with those of peers using the same or similar equipment/methodology.

**Precision** - Closeness of agreement between independent, repeated results obtained under specific conditions. These may be derived in the same day (intraday) or on different days (between or interday).<sup>1</sup>

**Proficiency testing (PT)** - A measure of laboratory competence during which an interlaboratory test comparison is conducted for the purpose of determining the laboratory's capability to conduct a specific diagnostic test. <sup>11</sup> PT is often used synonymously with EQA but may specifically refer to testing performed in compliance with state or federal regulations.

**Result (of a measurement)** - Value attributed to a measurand, obtained by measurement. A complete statement of the result of a measurement includes information about the uncertainty of measurement. Can be used synonymously with *value*.<sup>12</sup>

**Sample** - The appropriate representative part of a specimen which is used in the analysis.<sup>13</sup>

**Specimen** - Material available for analysis.<sup>13</sup>

**Standard deviation (SD, s)** - A measure of variability or diversity. It shows how much variation or dispersion there is from the average (mean or expected value) during repeated measures. A small SD indicates that the data points tend to be very close to the mean, whereas large SD indicates that the data points are spread out over a wide range of values. SD is the square root of a dataset's variance.<sup>1,14</sup>

**Standard deviation index (SDI, % error)** - A comparison between an individual laboratory mean and the peer group mean that reflects the systematic error or bias of a method. It may be a positive or negative number.  $SDI = (\text{Lab mean} - \text{Group mean}) / \text{Group SD}$ <sup>15</sup>

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{(n - 1)}}$$

**Total error (total analytical error)** - The sum of random error (imprecision) and systematic error (bias).

**Total error, allowable (TEa)** - a.k.a. desirable total error. An analytical quality requirement that sets a limit for the imprecision (random error) and inaccuracy (systematic error or bias) that are tolerable in a single measurement or single test result to ensure clinical usefulness. Per ASVCP guidelines, if expressed in units of %,  $TE_{obs} = 2CV + \text{bias}\%$ ; if expressed in analyte units,  $TE_{obs} = 2SD + \text{mean difference}$ .<sup>1</sup>

**Value (of a quantity)** - Magnitude of a particular quantity generally expressed as a unit of measurement which may be multiplied by a number (e.g.  $3.5 \times 10^3$  cells/ $\mu\text{L}$ , 5 IU/L).<sup>12</sup> It can be used synonymously with *result*.<sup>6</sup>

**Variable** - A quantity of interest, whose value or magnitude fluctuates or changes (e.g. creatinine).<sup>13</sup> It can be used synonymously with *analyte*.<sup>6</sup>

## 4. External quality assessment

### 4.1 Overview

4.1.1 One or more test materials are distributed to participating laboratories that submit their results to the provider. The raw data are analyzed, and the performance of each laboratory is reported in comparison to an appropriate peer group or to a known mean of a reference material.<sup>15</sup> Most providers of veterinary EQA offer results electronically and results can be received in approximately 4 to 6 weeks.

4.1.2 At this time, most currently available veterinary EQA programs are based on peer comparison testing, in which participating laboratories compare their performance to that of other participating laboratories (“peers”) using data compiled by the program provider. Human, and fewer veterinary, EQA programs may compare performance to known results for certified reference materials instead of peer results.

4.1.3 EQA test materials may be from a variety of sources and may contain additives or chemicals, as well as other materials inherent to the source of the specimens. Commercially available reference materials used in EQA may be human-derived, animal-derived, or both (e.g., serum from one species with added components from another species). Commutable materials are available for some analytes but are not common.

4.1.4 For comparisons to be most meaningful, the same EQA test material lot numbers, methodologies, and instruments should be used by the laboratories involved.

4.1.5 The format of the report and the type of metrics reported for the various analyses and serial evaluations may differ, depending on the EQA program and its design. **(See appendix C for an example of an EQA data and interpretation)**

4.1.6 Some programs include graphical representation of results; total error may be presented this way. A Youden plot is another depiction of lab performance. It is a two-

sample (control material) plot in which more than one laboratory can be represented. Vertical and horizontal lines or boxes within the plot denote ideal performance. More information on Youden plots can be found in the provided references.<sup>16</sup>

4.1.7 Results may be reported in a dichotomous scale (e.g. positive/negative; agreed/did not agree with expected results) for immunologic assays or morphologic evaluations in the veterinary field.

4.1.8 Educational tools and materials, such as newsletters and/or relevant articles, may also be provided by the EQA program administrator.

4.1.9 Participation in EQA programs is generally voluntary and utilized primarily to improve quality. Veterinary diagnostic laboratories could be required to participate in PT programs if they implement testing for certain infectious diseases or tainting compounds that are under surveillance and/or regulation by state or federal governments (e.g., testing for equine infectious anemia [Coggins test], scrapie, foot-and-mouth disease, and drug/chemical residue screening). Detailed discussion of PT for regulatory purposes is beyond the scope of this document. Please refer to the appropriate state or federal regulatory agency for more information.

## **4.2 Benefits of EQA participation**

4.2.1 Comparing results of one laboratory to those of others allows for greater monitoring of accuracy on a regular basis and greater oversight of quality and desired level of performance. For ideal assessment, identical methods and instruments should be grouped and compared.

4.2.2 Sample commutability impacts EQA. There is scarcity of matrix-free testing materials for most measurands evaluated in veterinary medicine, and test items may not be commutable across the various instruments and methodologies used by EQA subscribers. Carefully selected peer groups are theoretically subject to the same matrix effects, minimizing the impact of non-commutability on results.<sup>5</sup> Use of a human-derived matrix-free material assayed by a definitive or reference method has been included in human educational EQA assessments. Use of matrix free material removes the issue of non-commutability, simplifying comparisons. Although it would be ideal, cost and effort of developing materials commutable across methods and species for use in veterinary EQA may be cost prohibitive.

4.2.3 Potentially, instrument manufacturers or suppliers that provide or are associated with EQA programs may use the information obtained to internally monitor performance of their technology and implement necessary improvements.

4.2.4 Educational tools may be offered that can positively impact personnel knowledge and performance.

### **4.3 Limitations of EQA:**

4.3.1 Owing to differences between testing materials and patient samples and between EQA worksheets and patient report formats, it is not always possible for EQA events to reliably assess proficiency of sample preparation (i.e., detect pre-analytical errors) and results reporting (post-analytical errors) for the participating laboratories. Laboratories should establish other procedures for the evaluation of these potential sources of error, such as performing internal audits and/or practice runs and distributing questionnaires that address quality of reporting.<sup>4</sup>

4.3.2 When results are analyzed in relation to group mean and other data, the laboratory performance evaluation is only as good as the laboratories participating in the program. This limitation may be avoided by using a reference material as the test material and comparing laboratory results to the known (assayed) results for that reference material. If a commutable reference material is used, laboratory results may be compared to results from a definitive or reference method.

4.3.3 Combining more than one instrument and/or method type into a peer group is not recommended. Peer groups composed of heterogeneous instruments/methods are likely to have increased variation in the data compared to homogenous peer groups for the same measurands, and data may be insensitive for detecting diagnostically relevant errors. For certain laboratories using instruments and methodology less commonly available, peer groups may be deemed too small for practical comparison. Peer groups, therefore, may have to be selected based on the methodology or instrument alone instead of both. In choosing if a comparison across methods or instruments should be prioritized, the evaluator must have extensive knowledge of both the methodology and instrumentation in order to determine the most useful comparison for the individual laboratory's performance.

### **4.4 Recommendations for participating in EQA programs**

4.4.1 ASVCP recommends that both commercial and in-clinic laboratories participate in an EQA program.<sup>2</sup>

4.4.2 EQA is by nature a retrospective evaluation, as results are not generally available for 4 to 6 weeks after a test event. EQA programs are not meant to replace already implemented QA protocols. EQA participation should be used in addition to daily internal (meaning internal to the laboratory) QA/QC.

4.4.3 EQA program participation for hematology and biochemistry testing is considered to be of highest importance. A quality monitoring plan should be in place for all tests within the laboratory. For analytes not assessed by EQA, periodic comparison with

results from another laboratory should be considered. Readers are referred to section 5 for additional information on comparative testing between laboratories.

4.4.4 EQA testing is recommended quarterly. Combination with more frequent comparative testing (see section 5) may be necessary for instruments whose performance has been erratic or is unknown.

4.4.5 Archiving data and accompanying reports for at least 7 years is recommended. Archives may be paper or electronic. Electronic data must be appropriately secured and backed up.

4.4.6 EQA peer group composition and size, program participation costs, and testing materials used by the program should be known and understood by subscribing laboratory personnel. Request this information from the provider before choosing the appropriate EQA program.

4.4.6.1 Peer groups should use the same reagents, instruments, and analytical methods.

4.4.6.2 It is preferable to participate in programs whose eligibility and/or cost does not hinge on the purchase and continued use of a particular instrument and control materials or reagents. Similarly, the EQA provider should be independent of the instrument manufacturer.

4.4.6.3 Testing materials that arrive in liquid form are preferred for in-clinic laboratory EQA, as they do not require reconstitution (and thus eliminate potential measurement errors associated with pipetting and mixing). On occasion, lyophilized test materials are necessary, largely due to concern for sample deterioration. Test materials requiring reconstitution leave more room for preanalytical error than premixed test items, which can adversely affect outcome.

**See Appendix A: Summary of recommendations for EQA participation**

**See Appendix B: List of vendors/manufacturers supplying EQA services**

#### **4.5 Handling and processing testing materials**

4.5.1 Upon receiving the test materials, laboratory staff should confirm that these arrived in a timely manner, within their reported shelf life, and in the intended condition. For example, if samples meant to arrive refrigerated are received at room temperature, the program provider should be notified, as results will likely be affected.

4.5.2 EQA provider instructions for sample handling should be strictly followed. Testing materials should not be modified from their original constitution by the participating lab unless instructed to do so by the EQA provider. Reconstitution and other modifications should be performed according to the provider's instructions. Contact the test material provider for assistance, as needed.

4.5.3 The EQA test materials should be processed identically to patient samples to the best of the staff's ability. All staff members normally responsible for instrument and patient sample handling should participate in EQA testing. Any personnel involved in sample accessioning or reporting should be included in the same process. The goal is to mimic day-to-day laboratory activities as much as possible.

4.5.4 Test items should be analyzed within a period of time that minimizes specimen deterioration. Results should be submitted in the units provided by the analyzing instrument.

#### **4.6 Evaluation of EQA performance**

4.6.1 Numerical and interpretive EQA reports should be assessed upon receipt from the program provider. They should also be archived so that EQA data may be monitored over time. Single unacceptable results may be sentinels of systemic laboratory problems, and reviewing historical data may aid trouble-shooting of unacceptable results and identifying potential causes. Typical acceptance criteria for individual laboratory results are values within +/- 2SD or +/- 3SD or +/- TEa of the peer group mean. Additional information regarding TEa can be found in the ASVCP guideline on allowable total error for biochemistry.<sup>1</sup> Calculation of SDI index may also be of use. An SDI of 0 is desired. An SDI of 1.0 or greater reflects a potential systematic error or bias that may lead to unacceptable results. An SDI >2.0 should be investigated regardless of the test.<sup>14,15</sup> The calculated value may be a positive or negative number, but interpretation of SDI should be based on the absolute value. Graphing calculated values over time may provide useful perspective regarding laboratory performance.

4.6.3 Some EQA providers offer continuing education programs and assistance with data interpretation. These are valuable resources that should be utilized whenever possible. See Appendix C: Example of an EQA data report and interpretation

#### **4.7 Managing unacceptable results**

4.7.1 All laboratory personnel should be informed of the results of laboratory EQA performance. One or more individuals should be clearly assigned tasks related to EQA performance analysis, including troubleshooting of any unacceptable results, implementation of corrective and/or preventative actions, and monitoring and documentation of these measures in order to ensure that unacceptable performance has been resolved. Procedural and/or policy changes resulting from EQA performance

analysis must be completely documented and transparent to laboratory personnel and clients.

4.7.2 Problems identified should be classified into specific categories whenever possible, recognizing that each individual laboratory has its own documentation system based on its overarching quality management system. CLSI suggests the following error categories:<sup>4</sup>

- Clerical errors
- Methodological problem
- Equipment problem
- Technical errors
- Problems with EQA testing material
- Problem with evaluation of results
- No explanation after investigation

The “no explanation” category should be reserved for cases in which extensive research of the problem has not revealed a likely cause. In those situations, if subsequent EQA events indicate acceptable performance, then the original problem may have been due to random error. If the unacceptable results persist in subsequent EQA events, a systematic error is likely and additional steps may be needed to ensure the quality of laboratory results (See section 5 -- Comparative Testing). More information on random and systematic error may be found elsewhere.<sup>1,17</sup> Classifying and documenting the problem as specifically as possible allows for identification of the root cause. For example, a problem with a test result may be identified whose root cause may be SOPs that are outdated, unclear, or not reviewed by personnel.

4.7.3 If problems with a test material are identified, and inadequate handling or storage by the participating lab has been excluded, this information should be communicated to the EQA provider prior to running the testing materials. This step may help the lab avoid processing of an inadequate test material aliquot and also provides feedback to the EQA provider.

4.7.4 If an unacceptable EQA result is obtained, consideration should be given to clinical importance of the particular analyte and concentration. For example, if a lab has unacceptable performance (based on statistical or other criteria) for an analyte concentration that is not medically relevant or likely to alter patient management (e.g., low ALP activity), then trouble-shooting (while prudent in the long term) is prioritized lower than investigating unacceptable results for a more medically relevant analyte and concentration (e.g., hypoglycemia). Use of ASVCP allowable total error guidelines for biochemistry may be of value in assessing the “unacceptability” of results.<sup>1</sup>

4.7.5 When difficulties arise and/or unacceptable results are obtained, consider contacting the instrument manufacturer’s customer support staff, reagent

manufacturer, veterinary clinical pathologist, or other knowledgeable professional in order to identify problems and develop a corrective action plan.

## 5. Other Comparative Testing

### 5.1 Other comparative testing -- Overview

5.1.1 Some university, commercial reference, or in-clinic laboratories may have two or more instruments measuring the same analytes, either with the same or differing methodologies. If the same reference intervals are to be utilized across multiple analyzers, it is critical that the analyzers have similar analytical performance and that reference intervals are validated for use with all instruments. For additional information about reference interval transfer and validation, see “ASVCP Guidelines: determination of de novo reference intervals in veterinary species and other related topics”.<sup>18</sup>

5.1.2 Point-of-care testing (POCT) sites (including private veterinary practices and individual patient wards within larger hospitals) may compare test results from POCT instruments to those from a university laboratory or a commercial or reference laboratory. Such comparisons may be done to verify POCT findings and/or monitor the accuracy and reliability of POCT results. Defined acceptance criteria should be established in advance, in order to decide whether the results obtained by a POCT instrument and a reference/commercial/university laboratory instrument are comparable. Acceptance criteria may be based on (but are not limited to) +/- total error (TEa), +/- 2SD or +/- 3SD, or degree of difference from reference interval of each instrument.

5.1.3 Either within or across laboratory systems, results of comparative testing are often used to determine bias and help laboratorians understand how results generated by one instrument compare with those generated on a different instrument (e.g. “the sodium on this instrument runs high”).

### 5.2 Frequency and design of comparative testing

5.2.1 Frequency and design of comparative testing events will vary depending on the testing site’s needs and quality management system. The need for comparative testing and the number of analytes tested depends on historical instrument/method performance (i.e., how stable is the method?), concurrent EQA participation, history of suspicious results, and laboratory client or instrument operator feedback. For these reasons, only general recommendations are given below. Tailored recommendations can be sought via consultation with a veterinary clinical pathologist and/or a statistician skilled in quality management.

5.2.2 Comparative testing should be considered:

5.2.2.1 *Based on special causes.* Special causes that may trigger further investigation of analytical performance via comparative testing include unexpected result(s) based on clinical findings, changes in frequency or the nature of unexpected results detected in routine monitoring, unacceptable performance in an EQA event, unacceptable internal statistical QC performance, reagent or calibrator lot change, major maintenance or part change on an analyzer, or analyzer software update.<sup>9</sup> It is worth noting that this kind of comparative testing is not proactive. Rather, is a post hoc analysis, occurring after problems have already arisen. Comparative testing based on special causes may be used to verify the accuracy of a result based on predefined acceptance criteria and determine if further investigation or troubleshooting of instrument performance should be undertaken.

5.2.2.2 *Frequently (e.g., weekly).* Frequent testing is recommended for tests/methods whose control material may not reflect all analyte concentrations of medical importance (e.g. glucose, which has many potential clinical decision levels ranging from profound hypo- to hyperglycemia). For such analytes, patient samples or “spiked” samples (see section 5.4.1.2) with extreme high or low results may be compared with those of an outside laboratory to ensure acceptability with predefined acceptance criteria. See ASVCP allowable total error guidelines for biochemistry for assistance determining result acceptability.<sup>1</sup>

5.2.2.3 *Quarterly.* Periodic testing is appropriate for tests/methods not routinely offered in commercially available EQA programs. It may be helpful to coordinate the comparative testing with scheduled EQA test events.

### **5.3 Design of comparative testing events**

5.3.1 Single analyses of sample aliquots from a single specimen are often conducted to check suspicious patient results. This is an appropriate first step when investigating suspicious POCT results by comparing them to results from a university, commercial or reference laboratory. This design has significant limitations, however, and is not recommended for comprehensive investigation of suspicious patient results or POCT performance. A lack of comparability between paired, single analyses of specimens may not indicate a systematic problem but could simply reflect random error (imprecision) of either analyzer and/or bias between instruments.

5.3.2 More complex comparative study designs, such as utilizing multiple patient specimens covering a range of values and duplicate measurements of patient samples, should be used for comprehensive investigation. These designs will vary according to needs of the testing site and intended use of the comparability data.

5.3.3. When interpreting results from comparative testing, one instrument or method is designated as representing the analyte’s “true” value (e.g., a reference laboratory’s

instrument is used as the comparative method for a POCT instrument). In this scenario, any difference in results (e.g., any identified bias) is “assigned” to the instrument under evaluation (e.g., the POCT instrument). This is problematic if a large degree of bias exists between the two instruments being compared, or if precision of the two instruments varies. In that case, assigning all error to the instrument under evaluation unfairly overestimates error of that instrument and may be misleading. Some manufacturers offer suggestions for acceptable comparative instruments or methods. In all cases, the comparative instrument or method should be chosen carefully, and considerations for choosing a comparative method include:

- Purpose of the comparative testing event (how will the data be used?)
- Nature of the comparative method (definitive method, reference method, or field method)<sup>19,20</sup>
- Whether the comparative method is being used by a facility with known expertise in its use
- Whether the comparative method has similar analytical performance and calibration to the method being evaluated

In veterinary practices and many diagnostic laboratories, definitive methods, as defined by Tietz, are not available, and comparisons are made between reference and field methods or between two field methods.<sup>19</sup> Both methods will have some degree of inherent imprecision and bias, and comparative testing data can be interpreted as recommended in section 5.5. Reference laboratories to which samples for comparative testing are sent should be able to provide precision and total error information, which aids in comparison. For further review of method comparison and bias assessment, the reader is referred to other resources.<sup>20,21</sup>

## **5.4 Specimens used for comparative testing**

5.4.1 Individual patient specimens or pools of patient specimens are the preferred samples. Care should be taken to ensure that these are tested within a short period of time in accordance with the stability of the measurand(s) being evaluated. Patient specimens may not be adequate for comparison of hematologic analysis performed by laboratories some distance from each other due to the fragile nature of live cells in whole blood.

5.4.1.1 Patient samples should be of good quality and free of potential interferents or other potentially confounding factors (e.g. hemolysis, lipemia, icterus).

5.4.1.2 In some cases, spiking of an individual patient specimen or specimen pool may be required to obtain desirable levels of measurand(s) that cover all medically important levels. This is acceptable, but the material used to spike the specimen should be carefully considered (e.g., purified molecule of interest vs. patient sample containing that molecule as well as other compounds and matrix). This task is typically undertaken by larger laboratories with sufficiently knowledgeable staff and resources. Small or in-

clinic laboratories interested in this design should consult a clinical pathologist or other qualified professional for help with development and implementation.

5.4.2 Quality control materials, standards, calibrators, linearity verification materials, regulatory proficiency testing specimens, and EQA test materials may all be suitable for use in comparative testing, but care should be taken to determine the stability of such materials and ensure standardization of handling. It should also be determined if results obtained using these materials are commutable across all of the instruments/methods being compared and with patient specimens. Information about commutability is often provided by instrument and material manufacturers.

## 5.5 Basis for determining results comparability

5.5.1 There is no single method that is universally applicable for determining comparability of results. Several approaches exist and include:

### 5.5.1.1 Allowable total error recommendations

- TEa recommendations for biochemistry tests are available from the ASVCP, and their use is recommended.<sup>1</sup> Laboratories also may wish to establish their own quality specifications based on the species of interest, intended use of the results, client needs, or other factors. Developing quality specifications requires knowledge of clinical decision thresholds and analytical performance of the instruments in question. Dialogue between clinicians and laboratory management is needed for this process.
- Using TEa for determining comparability requires that one instrument is designated as the comparative instrument, and the result being compared (results from the index instrument) should fall within the range delimited by comparative instrument result +/- TEa.
  - Ex: The comparative result for albumin is 3.0 g/dL, which is within the laboratory's reference interval. The recommended TEa for albumin concentration within the reference interval is 15 %. Therefore, the index instrument result (result being compared) should fall within the range of  $3.0 \text{ g/dL} \pm 15\% = 3.0 \text{ g/dL} \pm 0.45 \text{ g/dL} = 2.6 \text{ g/dL} - 3.5 \text{ g/dL}$ .<sup>1</sup>

### 5.5.1.2 Reference intervals

Using reference intervals to define comparability of results is most useful when single results are obtained on a specimen from different instruments using different analytical methods. This is a situation that commonly occurs when POCT results are compared with those of university, reference or commercial laboratories. This approach also assumes that the reference intervals have been correctly established and truly reflect the biases inherent in each instrument/method. Placement within reference interval (upper half, upper

quarter, lower half, lower quarter) should be considered, as well as the degree of deviation below or above reference interval (between upper reference limit and 0.25 X upper reference limit, between 0.25 and 0.5 X upper reference limit, etc.). Results are acceptable if both fall within the same area of the respective reference intervals or deviate from the respective reference limits by the same magnitude.

#### 5.5.1.3 *Statistical criteria*

Statistical criteria that can be used to define comparability of two test results include SD and CV (or multiples thereof). SD and CV are derived from analytical performance data of the instruments in question. Again, one instrument is designated the comparative instrument, and the comparability criterion (e.g., results must fall within comparative instrument results  $\pm 3SD$ ) is derived from the comparative instrument's performance data. This approach likely requires advanced statistical knowledge in order to select the most appropriate statistical criterion. The methods used will depend on the design of the comparative study and intended use of the data.<sup>21,22</sup> The CLSI document EP31-A-1R contains additional information and examples that may be of benefit for laboratories wishing to conduct these types of statistical evaluations.<sup>9</sup>

#### 5.5.1.4 *Use of biologic variation (BV) data.*

This approach requires knowledge of biologic variation (BV) of the measurand in the species of interest, which limits application in veterinary medicine. Advanced statistical knowledge is also needed. Interested parties are referred to other publications for further information on this approach.<sup>23,24,25</sup>

**See appendix D: Example of a special cause comparability evaluation using reference intervals provided**

### **5.6 Clinical significance of result comparability**

5.6.1 The use of advanced statistical analyses yields the greatest probability of determining statistical power and statistical significance of comparative testing results. This may require consultation with a statistician or other knowledgeable professional. However, even the less complex evaluations outlined in section 5.5 can provide a perspective on clinical significance of results and information which helps ensure that accurate and reliable laboratory results are consistently obtained and produced.

5.6.2 Lack of comparability, particularly between paired, single analyses of specimen aliquots may not indicate a systematic analytical problem and could merely reflect random error (imprecision) or bias between instruments. Detailed investigation and possibly further testing are needed to determine if systematic or random analytical error is occurring, or whether lack of comparability could be due to pre-analytic or post-analytic factors.

## 5.7 Reasons for non-comparable results

5.7.1 Reasons for discrepancies in results include, but are not limited to:

- Differences in analytic methodology
- Differences in calibration between instruments/methods
- Differences in the precision between instruments/methods
- Lack of test material commutability (i.e., matrix effect)
- Simultaneous use of calibrator lots of differing ages and/or differing degradation in different laboratory locations
- On-instrument reagent degradation after calibration
- Instrument performance drift
- Use of different reagent lots or differences in packaging, transport and/or storage when the same method is used on more than one instrument
- Differences in instrument analytic parameters, such as dilution ratios and incubation times between different instruments that use the same reagents
- Pre-analytic effects on the sample, including differences in sample handling

5.7.2 Differences in calibration, reagent lots, and instrument parameters should be eliminated if possible. This can be addressed by standardizing calibrators, reagent lots, and analyzer settings in laboratories using multiple instruments to measure the same analytes whenever possible and applicable.

5.7.3 Differences in analytic methods cannot be minimized or eliminated if the instruments being compared are a reference laboratory instrument (e.g., laser-based hematology instrument) and a POCT instrument (e.g., an impedance counter for hematology testing). Bias between the instruments will be reflected in properly established reference intervals for each instrument, and interpreting comparative testing results as presented in section 5.5.1.4 is advised.

## 6. Acknowledgements

The authors would like to thank the members of the ASVCP's Quality and Laboratory Standards Committee for their thorough review of this manuscript.

## 7. References

1. Harr KE, Flatland B, Nabity M, Freeman KP. ASVCP guidelines: allowable total error guidelines for biochemistry. *Veterinary Clinical Pathology*. 2013; 42(4): 424-436. Available at: <https://www.asvcp.org/membersonly/pdf/TEa%20Guidelines%20Draft2.pdf>.
2. Flatland B, Freeman KP, Vap LM, Harr KE. ASVCP guidelines: quality assurance for point-of-care testing in veterinary medicine. *Veterinary Clinical Pathology*. 2013; 42(4): 405-423. Available at:

<http://www.asvcp.org/pubs/qas/newQas/PDF/ASVCP%20POCT%20QA%20Guideline%20May%202013.FINAL.pdf>.

3. Westgard JO. *Basic QC Practices*. 3<sup>rd</sup> ed. Madison, WI: Westgard QC, Inc.; 2010: 3-14.
4. Clinical and Laboratory Standards Institute (CLSI). Using proficiency testing to improve the clinical laboratory; approved guidelines, 2<sup>nd</sup> ed. (GP27-A2). Wayne, PA: AALA; 2007.
5. Miller WG. The role of proficiency testing in achieving standardization and harmonization between laboratories. *Clinical Biochemistry*. 2009; 42: 232-235.
6. Braun JP. Communicating with precision in veterinary clinical pathology: definitions, units, and nomenclature. *Vet Clin Path* 2009; 38(4): 416-417.
7. Fraser CG. *Biologic Variation: From Principles to Practice*. Washington, DC: AACC Press; 2001: 1-27.
8. Armbruster D, Miller RR. The Joint Committee for Traceability in Laboratory Medicine (JCTLM): A global approach to promote the standardization of clinical laboratory test results. *Clin Biochem Rev*. 2007; 28: 105-113.
9. Clinical and Laboratory Standards Institute (CLSI). Verification of comparability of patient results within one health care system approved guideline. (EP31A-1R). Wayne, PA: AALA; 2013.
10. Laboratory Accreditation Bureau. Information. <http://l-a-b.com/information/traceability/>. Accessed March 31, 2015.
11. OIE web. Manual of Diagnostic Tests and Vaccines for Terrestrial Animals 2014. [http://www.oie.int/fileadmin/Home/eng/Health\\_standards/tahm/0.04\\_GLOSSARY.pdf](http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/0.04_GLOSSARY.pdf). Accessed April 17, 2015.
12. Westgard QC. Glossary of ISO terms. <http://www.westgard.com/isoglossary.htm>. Accessed March 31, 2015.
13. Westgard QC. WQC glossary. <http://www.westgard.com/glossary.htm>. Accessed March 31, 2015.
14. Westgard JO. *Basic QC Practices*. 3<sup>rd</sup> ed. Madison, WI: Westgard QC, Inc.; 2010: 175-194.
15. Westgard JO. *Basic QC Practices*. 3<sup>rd</sup> ed. Madison, WI: Westgard QC, Inc.; 2010: 195-212.
16. Medcalc web. Youden Plots. <http://www.medcalc.org/manual/youdenplot.php>. Accessed March 31, 2015.
17. Westgard JO. *Basic QC Practices*. 3<sup>rd</sup> ed. Madison, WI: Westgard QC, Inc.; 2010: 202-3, 233.

18. Friedrichs KR, et al. ASCVP reference interval guidelines: determination of de novo reference intervals in veterinary species and other related topics. *Vet Clin Path* 2012; 41(4): 441-453.
19. Tietz NW. A model for a comprehensive measurement system in clinical chemistry. *Clin Chem*. 1979; 25:833–9.
20. Flatland B, Friedrichs KR, Klenner S. Differentiating between analytical and diagnostic performance evaluation with a focus on the method comparison study and identification of bias. *Vet Clin Pathol*. 2014; 43(4): 475-486.
21. Jensen, AL and Kjelgaard-Hansen, M. Method comparison in the clinical laboratory. *Vet Clin Pathol*. 2006; 35(3): 276-286.
22. White, GH. Basic of Estimating Measurement Uncertainty. *Clin Biochem Rev* vol 29 suppl (i). 2008; S: 53-S60.
23. Petersen PH, Fraser CG, Westgard JO, Larsen ML. Analytical goal-setting for monitoring patients when two analytical methods are used. *Clin Chem*. 1992; 38: 2256-2260.
24. Harris EK. Statistical principles underlying analytical goal setting in clinical chemistry. *Am J Clin Pathol*. 1979; 72: 374-382.
25. Fraser CG, Petersen PH. The importance of imprecision. *Ann Clin Biochem*. 1991; 28: 207-211.

## Appendix A

### Summary of recommendations for EQA participation

Criteria	Comments
<b>Participation</b>	A combination of external quality assessment and use of comparative testing in independent laboratory quality assessment based on events defined by individual laboratories. Each laboratory should determine the combination that best fits their situation and the circumstances encountered in their laboratory
<b>Types of testing</b>	Hematology and biochemistry, additional testing as available and pertinent to the participating laboratory
<b>Peer group</b>	Made up of laboratories using same instruments, methods, and test material lot numbers
<b>Testing materials</b>	Liquid materials are preferable for in-clinic laboratory EQA, as these do not require reconstitution. Assayed materials are preferred.
<b>Frequency of EQA events</b>	Quarterly
<b>Submission of reports</b>	Electronic is preferred over manual
<b>Timeliness of reports</b>	Turnaround time of 8 weeks or less
<b>Evaluation of reports</b>	<ul style="list-style-type: none"> <li>• As soon as possible to minimize the potential effects on patient results</li> <li>• All staff should be informed of results</li> <li>• Results of comparative testing should be assessed using predetermined criteria, which may include +/- 2SD, +/- 3 SD, +/- TEa or other specified criteria</li> <li>• Rapid corrective action(s) should be taken to address any problem(s) identified with follow-up monitoring implemented</li> <li>• Investigation and corrective action prioritized according to medical importance of the analyte and analyte concentration</li> </ul>
<b>Services provided</b>	Educational content serves to improve laboratory practices When needed, consultation with clinical pathologist or other knowledgeable professional should be available
<b>Data archiving</b>	Seven (7) years
<b>Cost</b>	Should not hinge on the purchase and continued use of a particular instrument, control materials, or reagents
<b>Provider</b>	Should be independent of instrument manufacturer

## Appendix B

### **List of vendors/manufacturers supplying veterinary quality control materials and services**

Vendors and manufacturers are listed alphabetically. The ASVCP does not endorse any particular vendor or manufacturer. This list is for informational purposes only and does not constitute a legal contract or endorsement between ASVCP and any person or entity unless otherwise specified. Information on the ASVCP web site is subject to change without prior notice. Although every reasonable effort is made to present current and accurate information, ASVCP makes no guarantees of any kind.

#### Beckman Coulter (Electronic Quality Assurance Programs)

Diagnostics Division Headquarters  
250 South Kraemer Boulevard  
Brea CA 92821-6232  
Toll Free 800-526-3821  
Phone: 714-993-5321  
Fax: 800-232-3828  
<http://www.beckmancoulter.com/qap>

#### Bio-Rad Laboratories

Clinical Diagnostics Group  
4000 Alfred Nobel Drive  
Hercules, CA 94547  
Toll Free: 800-2-BIORAD (800-224-6723)  
Fax: 510-741-6373  
Email: [diagcs@bio-rad.com](mailto:diagcs@bio-rad.com)  
<http://www.bio-rad.com/en-us/category/external-quality-assurance-services-eqas>

#### College of American Pathologists

325 Waukegan Road  
Northfield, IL 60093-2750  
Toll Free: 800-323-4040  
Phone: 847-832-7000  
Fax: 847-832-8000  
<http://www.cap.org/web/home/lab/proficiency-testing>

#### European Veterinary Endocrine Quality Assurance Scheme (EVE-QAS)

Veilingweg 1a  
4731CW Oudenbosch  
The Netherlands  
Phone: 31 (0) 165 504488  
Fax: 31 (0) 165 504711  
[endocrinevet.com/eve-qas.html](http://endocrinevet.com/eve-qas.html)

Eurotrol B.V.  
P.O. Box 722  
6710 BS Ede  
The Netherlands  
Phone: +31 318 695777  
Fax: +31 318 695770  
<https://www.eurotrol.com>

Eurotrol, Inc.  
One Wall Street  
Burlington, MA 01803  
Toll Free: 866-234-5754  
Phone: 978-598-3779  
Fax: 978-598-3780  
Email: [officeUSA@eurotrol.com](mailto:officeUSA@eurotrol.com)  
<https://www.eurotrol.com>

Insight<sup>™</sup> Interlaboratory Quality Assessment Program from Sysmex  
One Nelson C. White Parkway  
Mundelein, IL 60060  
Toll Free: 888-879-7639 (USA)  
[insight@sysmex.com](mailto:insight@sysmex.com)  
<https://www.sysmex.com/us/en/Pages/Registration.aspx>

National Institute of Standards and Technology  
100 Bureau Drive, Stop 1070  
Gaithersburg, MD 20899-1070  
Phone: 301-975-NIST (6478)  
[inquiries@nist.gov](mailto:inquiries@nist.gov)  
<http://www.nist.gov/mml/csd/qaps.cfm>

NEQAS  
UK NEQAS Office  
PO Box 401  
Sheffield  
S5 7YZ  
<http://www.ukneqas.org.uk/content/Pageserver.asp>  
Phone: 0114 261 1689  
Fax: 0114 261 1049  
email: [office@ukneqas.org.uk](mailto:office@ukneqas.org.uk)  
<http://www.ukneqas.org.uk/>  
Randox Laboratories (RIQAS)  
Headquarters (Ireland)  
55 Diamond Road, Crumlin

Co. Antrim  
United Kingdom  
BT29 4QY  
Phone: +44 (0)28 9442 2413  
Fax: +44 (0)28 9445 2912  
[www.randox.com](http://www.randox.com)

Randox Laboratories-US, Ltd.  
515 Industrial Boulevard  
Bardane Industrial Park  
Kearneysville, WV, 25430  
Phone: 866-4-RANDOX  
Fax: 866-RANDOX 1  
<http://www.randox.com>

URIKA Pathology, LLC  
8712 53<sup>rd</sup> Place W  
Mukilteo, WA 98275  
Phone: 352-258-4055  
[info@urikapathology.com](mailto:info@urikapathology.com)  
[www.urikapathology.com](http://www.urikapathology.com)

Veterinary Laboratory Association (VLA) Quality Assurance Program  
University of Prince Edward Island  
550 University Avenue  
Charlottetown, PEI  
Canada C1A 4P3  
Phone: 207-227-0302  
Fax: 207-433-1018  
[qap-avc@upei.ca](mailto:qap-avc@upei.ca)  
[vla.timelessveterinarysystems.com](http://vla.timelessveterinarysystems.com)

## Appendix C

### Example of an EQA data and interpretation

Component	Sample	Last Evaluated	Units	
Sodium	2013-1A-CHEM- Canine	2013-03-06 05:47:14	mmol/L	
Your result	145.00 mmol/L			
Description	N	Mean	2 X SD <sup>a</sup>	CV%
All Results	175	139.94	5.35	1.9
Your Method ISE- diluted	121	139.89	5.02	1.8
Your Instrument Hitachi 902-917/Mod P	7	140.43	4.12	1.5

<sup>a</sup>2 X SD: two (2) standard deviations

The table shows results for Sodium (Na). It depicts a comparison of laboratories using various instruments as well as laboratories using similar methodologies and the same instrument as the submitting laboratory. The number of results (N) in each group varies.

- The analyte is identified as Sodium; the unit of measurement is mmol/L
- The testing material is identified as 2013-1A-CHEM-Canine and monthly statistics are for March
  - Note: Comparisons are best performed utilizing data from the peer group with the same instrument and methodology. To the authors' knowledge, recommendations regarding minimum peer group size do not exist, as numerous factors contribute, including biologic variation and clinical significance of an abnormal result. In this example, the peer group for laboratories using the submitting laboratory's instrumentation is relatively small. Therefore, while not ideal, it is reasonable that interpretation initially is focusing on the peer group using the submitting laboratory's methodology rather than the peer group using the same instrument.
- The monthly peer group mean for laboratories using all methods, same methodology, and same instrument are listed along with their respective CV in the above table. Note that 2 X SD and CV% grow progressively smaller as the comparison becomes specific to the same method and instrument.
- The measured value for the analyte as performed by the submitting laboratory is 145.00 mmol/L. For all comparisons, the submitting laboratory's result is of higher numerical

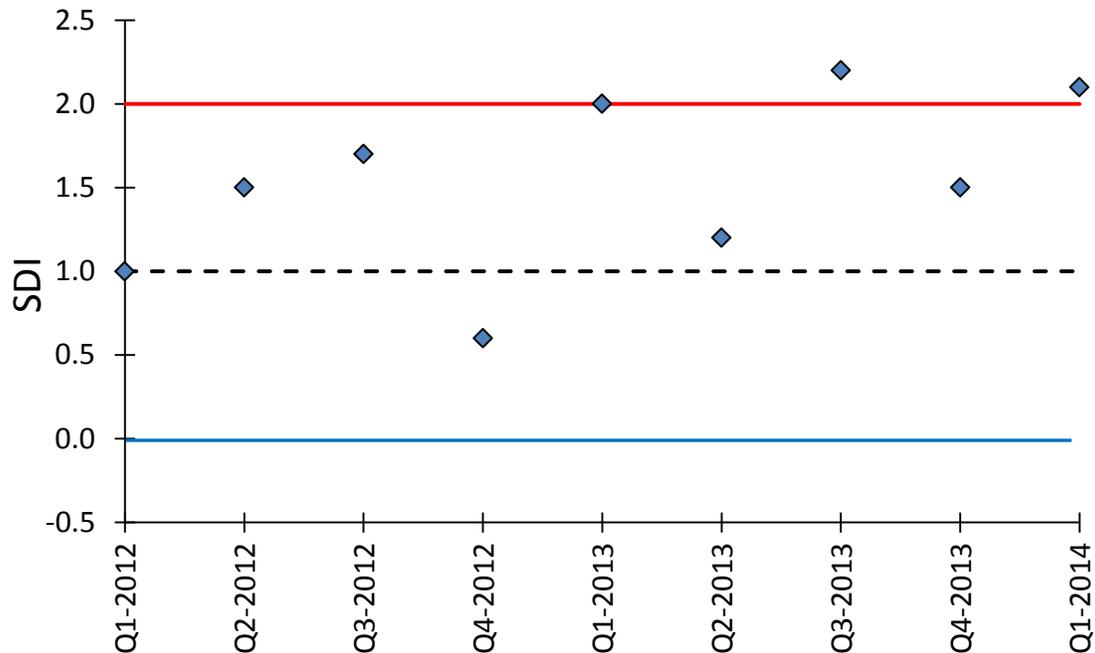
value than the peer group mean. Acceptance limits for the measured value are calculated in three ways below. In case of regulatory or accreditation requirements, acceptance limits should be calculated accordingly.

- Peer group mean +/- 2SD. The measured value is acceptable relative to the “all results” mean. However, the measured value falls outside of 2SD from both the specific instrument (140.43 +/- 4.12 or 136.31-144.55) and specific method (139.89 +/-5.02 or 134.87-144.91) group means. The result is unacceptable relative to those two group means.
  - Peer group mean +/- 3SD. The measured value is acceptable compared to all group means (e.g., 145.00 is within the range delimited by 139.89 +/- 3SD, or 139.89 +/- 7.53).
  - Peer group mean +/- TEa. The measured value is acceptable compared to all group ISE means using the ASVCP-recommended TEa for sodium of 5% (e.g., 145.00 is within the ranged delimited by 139.89 +/- 5%, or 139.89 +/- 6.99 = 132.9 – 146.88).
- Another means of comparing performance is utilizing the Standard Deviation Index (SDI). It is calculated by the formula:

$$SDI = (\text{lab mean} - \text{Group mean}) / \text{Group SD}$$

The calculated value may be a positive or negative number, but interpretation of SDI should be based on the absolute value. Graphing calculated values over time (see below) may provide a useful perspective regarding laboratory performance. In this case, the standard deviation for the same methodology peer group is 2.51 (data not in table;  $5.02/2 = 1SD = 2.51$ ). The calculated SDI for the testing event is 2.0.

## Quarterly SDI Index



**Figure 1:** Graph representing the SDI index results for nine quarters of EQA testing for a laboratory. All values fall in the positive side of the scale. The blue line represents the desired SDI index (0). The dashed black line represents an SDI of 1.0; the solid red line represents an SDI of 2.0. The data point Q1-2013 represents SDI for the data depicted on the previous table. Serial SDI calculations reveal an upward trend in SDI index values. Points for the Q3-2013 and Q1-2014 events are above the red line, indicating a need for further investigation.

- After reviewing historical data, there is concern for a potential systematic error. Most of the values fall one standard deviation above the mean (e.g., SDI of 1.0 or greater) and at least two SDI calculations above 2.0 have been documented in more recent quarterly reports. The trend for this analyte should be closely observed and compared to previous and future months. Comparison of TEobs versus TEa may also be useful. The laboratory should consider evaluation of SOPs, reagents, and instrument performance analysis to identify potential sources of error, particularly if recent internal quality assurance results also suggest poor performance.

For quantitative results, plotting the standardized score (e.g., SDI, % error) on the vertical axis and the date of the testing event on the horizontal axis-- as shown in Figure 1-- is a simplified way of monitoring performance.<sup>4</sup> If results are consistently on one side of the zero line, a

systematic or calibration error is likely. Graphs which plot standardized scores vs. concentration over longer time periods may help demonstrate proportional bias (analyte concentration-dependent bias). These may be created using data provided and software within the laboratory information system utilized for quality assurance, when available.

## Appendix D

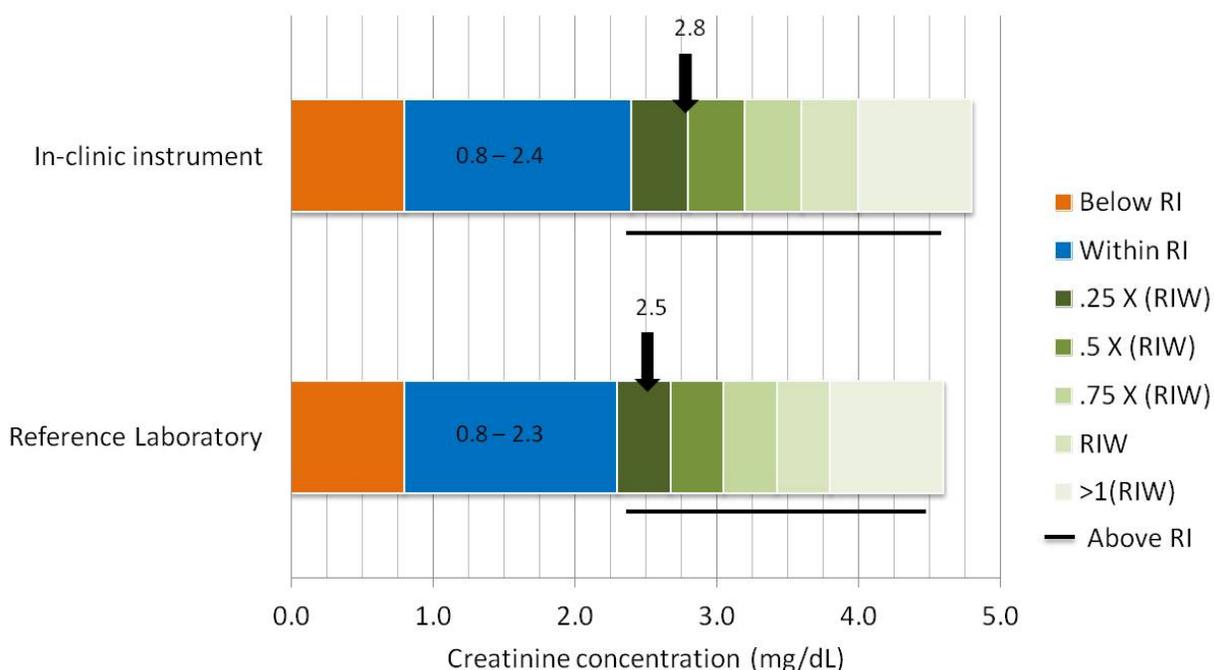
### Example of a special cause comparability evaluation using reference intervals

Analyte	Reference Laboratory		In-clinic instrument	
	Result (RI)	RIW	Result (RI)	RIW
Creatinine (mg/dL)	2.5 (.8 – 2.3)	1.5	2.8 (.8 – 2.4)	1.6
Albumin (g/dL)	2.5 (2.5 – 3.5)	1.0	1.9 (2.3 - 3.3)	1.0

RI: reference interval

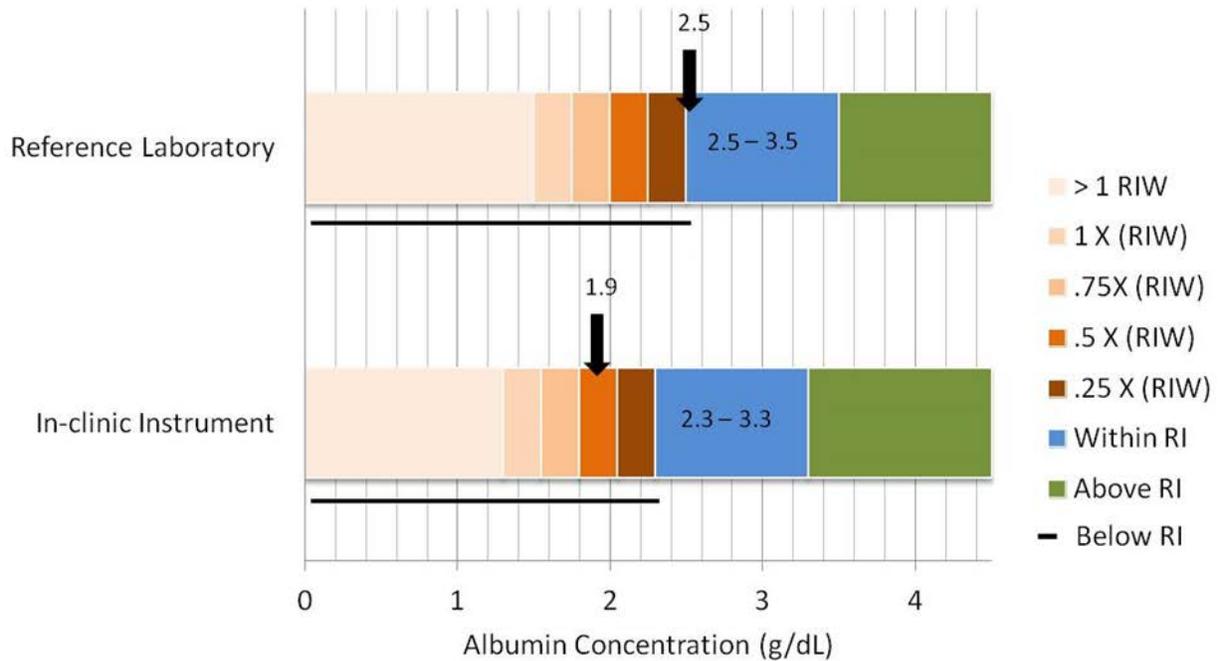
RIW: reference interval width = upper limit of RI – lower limit of RI

Serum from a patient was evaluated by two different biochemistry analyzers and the results are as above. The values reported by the reference laboratory and the in-clinic analyzer are different, but it is important to know whether they are comparable. Results can be compared using degree of deviation from the respective established reference intervals.



In this example, each result will be categorized by the degree of difference from the reference interval with respect to its width. This can be expressed in quarters or percentiles and calculated for each instrument [RI limit  $\pm$  (width of RI  $\times$  .25), for example]. The width of the creatinine reference interval is 1.5 units for the reference laboratory and 1.6 for the in-clinic instrument. The value obtained by the reference laboratory is above the upper limit of the reference interval by a factor of .25X the width of the reference interval (i.e. falls within 2.3-2.7). Similarly, the result obtained by the in-clinic instrument is also above the upper limit of the reference interval by a factor of .25X the width of the reference interval (falls within 2.4-2.8). In

other words, these results are outside of the reference interval by the same degree and thus considered acceptable, i.e., although the numerical values are different, they are considered comparable.



The results obtained for albumin appear different and would likely alter clinical decisions. The width of the RI for both the reference laboratory and the in-clinic instrument is 1 unit. The reference laboratory reported a value at the lower limit of the reference interval while the in-clinic instrument reported a value below RI. The latter is outside of the reference interval by a factor of .5X the width of the RI (between 1.8-2.3). These results are not comparable. If the reference laboratory result is considered “true”, this indicates that results from the in-clinic instrument are not acceptable. Evaluation of in-clinic instrument performance and inquiry into the performance of the reference instrument is warranted.

An alternate way to compare these results would be to identify the RI as the comparative method and use the interval comparative result +/- TEa to determine acceptability of the in-clinic instrument result. Whenever an unexpected result is obtained, it is prudent to evaluate that result in the context of the patient’s other clinical and laboratory findings.

Appendix E  
**Compliance Checklists**

**Section 4 External Quality Assessment**

Guideline Item	Compliant?	Additional Comment(s) by Auditor
4.4.2 Controls are run prior to testing, in accordance with laboratory policy	<input type="checkbox"/> Yes <input type="checkbox"/> No	
4.4.1, 4.4.2 A quality monitoring plan is in place including an EQA program that meets laboratory needs	<input type="checkbox"/> Yes <input type="checkbox"/> No	
4.4.4 EQA testing is performed quarterly or more often as needed	<input type="checkbox"/> Yes <input type="checkbox"/> No	
4.3.1. Mechanisms are in place to detect potential pre- and post-analytic errors	<input type="checkbox"/> Yes <input type="checkbox"/> No	
4.5.1,4.7.3 Quality of EQA test items is verified upon receipt and problematic findings are communicated to EQA provider	<input type="checkbox"/> Yes <input type="checkbox"/> No	
4.5.2 EQA provider instructions are strictly followed	<input type="checkbox"/> Yes <input type="checkbox"/> No	
4.5.3 EQA materials are processed identically to patient samples	<input type="checkbox"/> Yes <input type="checkbox"/> No	
4.5.4 EQA testing is performed in a timely manner	<input type="checkbox"/> Yes <input type="checkbox"/> No	
4.6.1 EQA data and interpretation are assessed upon receipt	<input type="checkbox"/> Yes <input type="checkbox"/> No	
	<input type="checkbox"/> Yes <input type="checkbox"/> No	
4.6.1, 4.7.1 EQA data, interpretation, and procedural and/or policy changes resulting from EQA analysis are archived on paper or electronically for 7+ years	<input type="checkbox"/> Yes <input type="checkbox"/> No	
4.6.3, 4.7.5 EQA provided continuing education and other assistance are utilized, when necessary	<input type="checkbox"/> Yes <input type="checkbox"/> No	
4.7.1 All laboratory personal are informed of EQA results	<input type="checkbox"/> Yes <input type="checkbox"/> No	
4.7.1 Designated person(s) analyze, troubleshoot, and monitor EQA performance	<input type="checkbox"/> Yes <input type="checkbox"/> No	
4.7.2 Problematic EQA results are categorized and root cause(s)	<input type="checkbox"/> Yes <input type="checkbox"/> No	

identified and monitored		
4.7.4 Troubleshooting of problematic results is prioritized based on clinical significance	<input type="checkbox"/> Yes <input type="checkbox"/> No	

**Section 5 Comparative Testing**

Guideline Item	Compliant?	Additional Comment(s) by Auditor
5.1.2 Have defined acceptance criteria for independent laboratory comparisons	<input type="checkbox"/> Yes <input type="checkbox"/> No	
5.2.2.1 Use comparability testing to investigate unexpected test results and as designated in the quality management system	<input type="checkbox"/> Yes <input type="checkbox"/> No	
5.2.2.3 Conduct comparison with external laboratories when EQA programs are not available for a specific test	<input type="checkbox"/> Yes <input type="checkbox"/> No	
5.4.1.1, 5.4.2, Use good quality, appropriate samples and test materials for comparison within the stability timeframe of the measurands being evaluated	<input type="checkbox"/> Yes <input type="checkbox"/> No	
5.4.1.2 Consult knowledgeable professional(s) prior to altering samples for comparison	<input type="checkbox"/> Yes <input type="checkbox"/> No	
5.5 Use appropriate methods for determining result comparability (with expert consultation, as needed)	<input type="checkbox"/> Yes <input type="checkbox"/> No	
5.6.2 Conduct further investigation if and when results are not comparable	<input type="checkbox"/> Yes <input type="checkbox"/> No	