Troubleshooting Guide for Proficiency Testing Data

Quantifying Deviations

Interpreting Deviations

Suggested Actions

Incorporation of Daily QC into Interpretation of PT Performance

www.cap.org

Troubleshooting Guide for Proficiency Testing Data

Overview

Analysis of proficiency testing results can reveal problems even before there is a proficiency testing failure. The CAP Proficiency Testing (PT) evaluations facilitate the recognition of potential problems with the graphical plot of the relative distance of results from targets, as percentages of allowed deviation. By recognizing patterns in these graphs that are unlikely to represent normal measurement variation, laboratory managers can initiate an investigation and, if necessary, corrective action, pre-empting a proficiency testing failure, or worse, an adverse patient event.

PT performance problems arise from both systematic and random sources of error. In this context, systematic error is characterized by consistent differences between participant results and target values, say, for example, when all results for an analyte lie on one side of the target value. Larger differences suggest a greater degree of systematic error. Problems due to random error are suggested by results that, on average, are close to the target value, but include some results showing large deviations on one or both sides of their target values.

Sometimes, by reviewing results from multiple mailings, performance trends that could lead to proficiency testing failures are easily recognized. In addition to characterizing patterns of measurement variability, the CAP graphical plots can identify trends that would be missed without reviewing multiple PT events together.

Quantifying Deviations from the Peer Group Target

The evaluation report lists normalized results as an SDI. The SDI is obtained by subtracting the group mean from your result and then dividing by the group SD. Monitoring rules based on SDIs have been shown to provide useful information for self-interpretation of proficiency test data.[1]

The evaluation report also includes a graphical summary using the relative distance of your results from the target. We refer to this distance as the *allowed deviation*. Typically, the range of acceptable results is the target +/- the PT allowable error. For example, this could be the group mean +/- 20%. To get the allowed deviation, the target value is subtracted from your result and the difference is divided by the PT allowable error. As a final step, this ratio is multiplied by 100 so that differences from the target value are on a percent scale ranging from -100 to +100. If results are beyond -100 or +100%, an “x” is printed at that limit indicating that results exceed the graphical limits. Monitoring rules based on the allowed deviations have also been shown to provide useful information for self-interpretation of proficiency test data.[2]

Interpreting Deviations from the Peer Group Target

Table 1 provides a summary of three performance rules to identify possible analytical problems and suggested actions based on SDI values. Once the likely source of the problem is identified, see the list of suggested actions to resolve the performance problem. Prior to investigating potential analytical problems, it is important to rule out clerical or specimen handling problems. These errors, or blunders, can be identified because they generally are far beyond the usual values with SDIs ranging from ±3.2 to ±10. Causes of blunders include mislabeling errors, misplacing specimens in an analyzer rack, calculation errors, inappropriate reagents or standards, neglect, or clerical errors. Blunders are some of the most common exceptions that are noted when evaluating proficiency testing results.[3,4]

Table 1. Guidelines for monitoring PT performance using diagnostic information from the SDI values reported on the PT evaluations

SDI Rule	Comments	Suggested Actions
At least one result exceeds ± 2 SDIs	Review results to rule out possible problems; identify possible errors from non-analytical sources for results with very large SDIs	See listing of suggested actions for evidence of systematic or random error
The average of your SDIs is > 1.5, or, if negative, < -1.5	Participant needs to calculate the average SDI; published studies confirm large average deviations can reveal potential problems	See listing of suggested actions for evidence of systematic error
The difference between the largest and smallest SDI is > 4	Published studies confirm large differences can reveal potential problems	See listing of suggested actions for evidence of random error

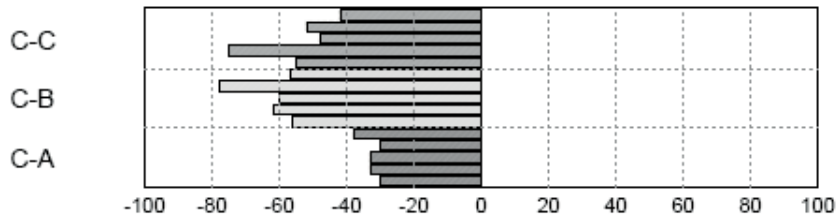
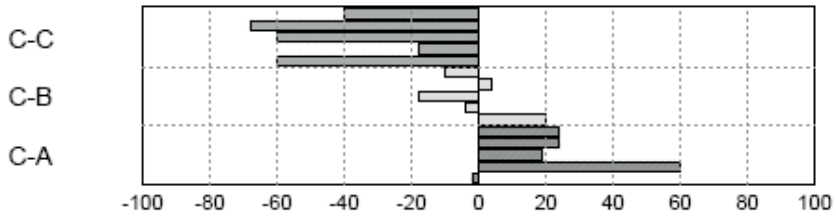
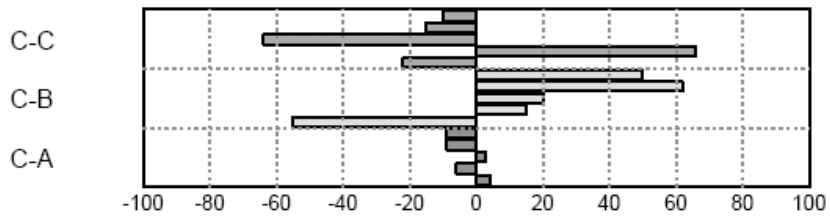
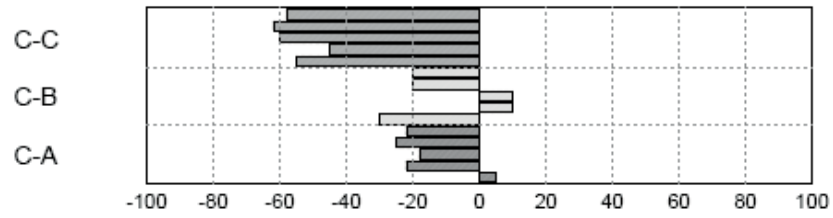
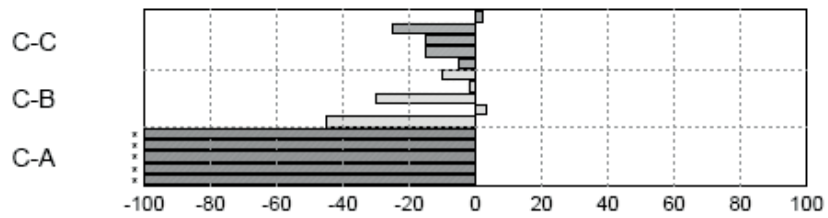
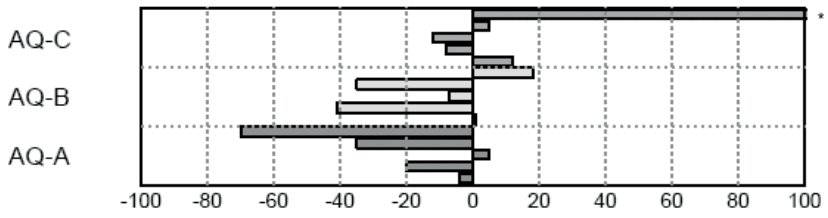
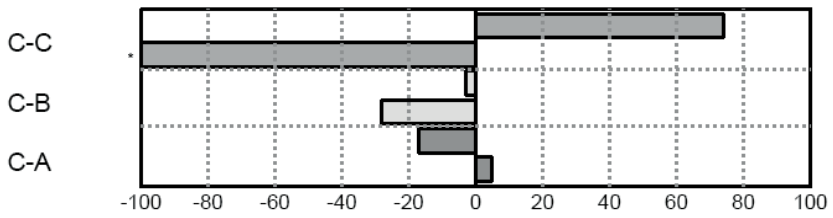
Table 2 provides performance rules based on the allowed deviations displayed in the graphical summaries. This table is divided into two sections reflecting interpretive guidelines for both single and multiple mailings. In some cases, identification of time-dependent trends can provide additional diagnostic information.

Table 2. Guidelines for monitoring PT performance using the evaluation graphs

Patterns in PT Evaluation Graphs for a Single Mailing		
Rule	**Comments**	**Suggested Actions**
One result in a mailing exceeds ±75% of the allowed deviation	Review results to rule out possible problems; identify possible errors from non-analytical sources such as clerical errors for results that exceed ±100% of the allowed deviation	If the data fail either of the rules below, follow the suggested actions for systematic or random error, as appropriate
All results are on one side of the target values with at least 1 difference exceeding ±50% of the allowed deviation	Shows bias indicating a possible calibration drift; there would be less concern if the relative differences were all close to 0	See listing of suggested actions for evidence of systematic error
Large positive and negative differences; combined lengths of longest positive and negative bars is > 140 out of total range of 200	Shows possible random error	See listing of suggested actions for evidence of random error
Time Trends in PT Evaluation Graphs over Multiple Mailings		
Rule	**Comments**	**Suggested Actions**
Persistent results on one side of the target values	Shows persistent bias, even if small; recalibration should have occurred within this time frame	See listing of suggested actions for evidence of systematic error
Results flip from one side of the target to the other	Shows impact of system and/or process changes; longer bars are of more concern	See listing of suggested actions for evidence of systematic error
Over time, length of bars increase	A sudden shift may show impact of system and/or process changes; may reveal new source of either systematic or random error	Follow the suggested actions for systematic or random error, as appropriate
Over time, length of bars decrease	Shows impact of system and/or process changes, particularly as a result of corrective action	Retain as documentation that corrective action has been successful

The figures in Table 3 on page 5 show examples of time-dependent trends. The final plot is from a non-regulated analyte with only two challenges per testing event. Even though there are limited data points with results from non-regulated analytes, the same general rules can be applied to identify and troubleshoot problematic testing results.

Table 3. Examples illustrating various patterns in cumulative PT results

Evidence of persistent bias spanning recalibration. Review process of setting QC target values; evaluate performance with assayed control material.	
Results flip from positive to negative bias. Review records to confirm system and/or process change. Follow suggested actions for systematic error.	
Over time, lengths of the bars increase on both sides of 0. For this pattern, follow suggested actions for random error.	
Over time, lengths of the bars increase primarily on one side. For this pattern, follow suggested actions for large systematic error.	
Lengths of the bars decrease. Corrective action following a previous failure can be easily demonstrated.	
Plot shows a result exceeding ±75% of the allowed deviation. This problem was due to a transcription error where results for Hgb and hematocrit were switched.	
Many PT challenges are for non-regulated analytes that can be identified as having only two samples. The same general patterns appear for non-regulated analytes, though with fewer data points on each plot. Here the C mailing samples were switched.	

Suggested Actions If There Is Evidence of Systematic Error:

1. Review internal quality control (QC) performance. Look for trends or shifts that may not yet trigger your rejection rules. Assess the process of setting and changing QC target values.
2. If recalibration has not already occurred, recalibrate the instrument.
3. If participating in an external QC performance program, review comparative reports for QC performance. If the laboratory performance on a lot of QC material is at consistent variance with the group performance mean, further investigation is warranted.
4. Use assayed control material to evaluate performance.

Suggested Actions If There Is Evidence of Random Error:

1. Rule out errors from non-analytical sources (transcription error, misplaced specimens, calculation error).
2. Investigate components of the analytical system (sample probes, reaction cells, reagents).
3. Review internal QC performance. Look for trends or shifts that may not yet trigger your rejection rules. Assess the process of setting and changing QC target values.
4. Use assayed control material to evaluate performance.

Additional Comments on Incorporating Daily QC into the Interpretation of PT Performance

When reviewing proficiency testing performance, it is important to identify current and potential failures by inspecting the SDIs and graphs of relative distances. Evaluation of your QC data preceding the challenge, at the time of the challenge, and following the challenge can also help identify possible problems and solutions. The QC records should indicate when recalibration and reagent lot changes occurred. All other laboratory records used in evaluating the proficiency samples and reporting the proficiency results should also be collected and examined when reviewing possible sources of problematic PT results.

References

1. Carey RN, Cembrowski GS, Garber CC, Zaki Z. Performance characteristics of several rules for self-interpretation of proficiency testing data. *Arch Pathol Lab Med*. 2005;129:997-1003.

2. Jenny RW, Jackson-Tarentino KY. Causes of Unsatisfactory performance in proficiency testing. *Clin Chem*. 2000;46: 89-99.

3. Grannis GF, Gruemer H-D, Lott JA, *et al*. Proficiency evaluation of clinical chemistry laboratories. *Clin Chem*. 1972;18:222-236.

4. Northam BE. Whither automation? *Ann Clin Biochem*. 1981;18:189-190.

325 Waukegan Road
Northfield, IL 60093-2750

Programs and Resources

Accreditation
Offers the "gold standard" for laboratory accreditation

Proficiency Testing
Ensures precision and confidence for your laboratory

CAP 15189℠
Recognizes a sustainable Quality Management System

Education
Serves as a leading resource for information and education in the laboratory

Efficiency/Quality Tools
Allows more time for patient care

Advocacy
Represents the interests of pathologists in the government and regulatory arenas and in the private sector

Membership
Provides valuable benefits and leadership for all laboratory professionals

SNOMED Terminology Solutions™ (STS)
Offers consultation and education services to transform health information

0880/pdf/1110/44195