Radioimmunoassay Quality Control and Troubleshooting

Helen H. Drew

The Johns Hopkins Hospital, Baltimore, Maryland

This is the third continuing education article in a four-part series on radioimmunoassay. After reading and studying this article, the reader should be able to: 1) describe the procedures used for immunoassay quality control; 2) state acceptable limits of results; and 3) discuss troubleshooting approaches.

Quality control (QC) is a vital function of every laboratory performing immunoassays and, when applied in a knowledgeable and sensible fashion, it can be a laboratory's most valuable tool. A properly designed QC program ensures that results are within acceptable limits of accuracy and precision. In simpler terms, a quality control program helps you answer the following important questions about your assay: Are today's results reliable? May the same assay be performed tomorrow using the same reagents with the same reliable results? In addition to answering these simple questions, quality control will help you unravel the mystery of *why* when your assay fails.

ACCURACY AND PRECISION COMPARISONS

Associated with any analytical measurement are certain amounts of error. These errors are generally of two kinds: (a) systematic errors which manifest themselves in the form of a bias; and (b) random errors which are reflected in poor reproducibility or imprecision (*1*). Bias, or accuracy, is represented by the difference between a true known value and the value obtained from replicate measurements (Fig. 1).

Systematic errors leading to bias are, in most cases, difficult to analyze because of the inability to know true value. One method of assessing systematic errors of an immunoassay procedure is to perform a recovery study which will give an indication of accuracy. Recovery studies were discussed in the second article in this radioimmunoassay series (*2*).

Random errors resulting in poor reproducibility and imprecision can be quantitated and controlled. The precision of any single measurement is governed by a number of factors. Specific examples include: (a) the characteristics of the sample (proper collection and storage); (b) the quality of the pipettes; (c) centrifuges and water baths used in the procedure; and (d) the care and expertise of the technologist performing the assay. Another influential factor is the counting errors in the measurement of radioactivity. These counting and experimen-

tal errors define the overall random error associated with a single measurement. Precision studies assess random errors and tell us how closely the same measurement can be repeated within the same assay (intra-assay precision) or between assays at different times (interassay precision). Therefore, the goal of a quality control program defines acceptable error limits without compromising the medical significance of the results.

QUALITY CONTROL STATISTICS

A QC program uses the fact that repeated measurements of a single sample fluctuate over time because of random errors or imprecision. It must be determined whether these fluctuations are within acceptable limits or reflect an analytical error requiring troubleshooting action. The distribution of these errors is assumed to be Gaussian, a curve of normal distribution. Quantification of this fluctuation of values is determined by calculating a mean, standard deviation (s.d.), and coefficient of variation (CV). The mean, which is the average of the series of measurements, is the central value around which the overall fluctuation of measurements is assessed. The variance, or degree to which the values deviate, is expressed by the s.d., which is used to define acceptable limits of performance.

Fluctuations about the mean by less than 2 s.d. are normal. Most laboratories consider deviation by more than 2 but less than 3 s.d. as an indication that a group of values may be out
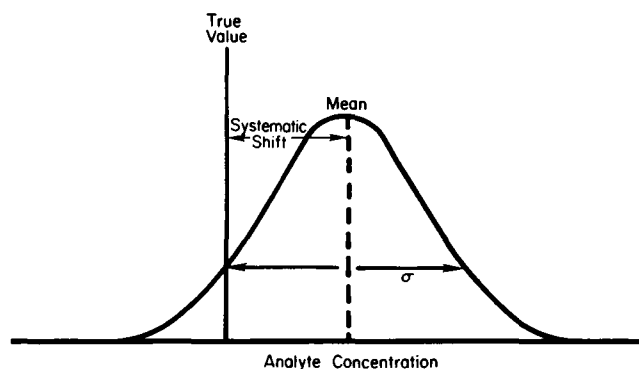

FIG. 1. Frequency distribution of multiple assayed results in a single sample. Systematic errors are indicated by the difference between a known true value and the assayed result. Random errors are indicated by the scatter of multiple results around the mean, represented by the standard deviation δ.

of control. Deviation by 3 s.d. indicates a problem and signals the beginning of troubleshooting action (Fig. 2).

Coefficient of variation is used as a precision indicator that expresses the s.d. as a percentage of the mean:.

$$\% \text{ CV} = \frac{\text{s.d.}}{\text{mean}} \times 100$$

A smaller %CV represents more precise and fluctuation-free measurements.

A good quality control program consists of establishing initial means, s.d.s, and CVs on control sera during the kit evaluation period. Following this criteria, a track record of the components of the assay and their performance over a period of time is kept. This track record is the basis for your daily decision to accept or reject results. Generally, 95% of all results will fall within ± 2 s.d. of the mean. Results outside 3 s.d. are statistically unlikely and signal a need to troubleshoot the assay. Furthermore, a shift or drift characterized by the presence of eight or more results on the same side of the mean indicates that troubleshooting action is needed. In addition, trends can be observed when a control continues to increase or decrease for six consecutive points (3). When this situation occurs, trends should be considered out of control and investigated.

A QC program should assess technique, assay conditions and equipment performance, and sample handling and overall quality of reagents on a regular basis. Each of these factors influence the overall quality of the results reported to the physician. As a starting point, examine your laboratory and its workload. Questions you should answer are: (1) Do technologists rotate assays or do the same individuals run the same assay regularly? (2) What tests are performed? (3) Are these screening tests or are results critical with narrow therapeutic ranges? (4) Examine the patient population on which the test will be performed—are they acutely ill hospitalized patients or outpatients? Laboratories performing therapeutic drug levels such as digoxin or the aminoglycosides need to strive for very low % CVs. Higher CVs may be acceptable for other tests. The answers to these questions will help you set sensible standards for performance in your laboratory and maintain the highest level of quality in your control procedures.

QUALITY CONTROL SAMPLES

The most commonly used indicator of an assay's performance in the laboratory is control material. A control material should simulate the physical characteristics and chemical composition of the specimen that is being analyzed. In addition, the matrix should be as close to the tested specimen as possible, so that it will act like an unknown when tested. In an assay, the control material is subjected to the same random and systematic errors as the patient's specimens.

There are several sources of control materials. Some controls are frequently supplied with kits by the manufacturer. However, these kit controls should never be relied upon exclusively because they are usually made up in the same matrix as the standards. Independent sources include pooled serum,

commercially assayed serum, and prepared in-house controls (4).

Serum pools can be easily made up from excess patient specimen material that has accumulated in the lab. Simply mix together all sera in the desired concentration range excluding icteric, lipemic, or hemolyzed sera. Mix thoroughly and filter to remove fibrin and other debris. Aliquot the remainder into individual tubes sufficient for one test, cover, and freeze. Allocate sufficient amounts of sera to last several months or as long as the analyte is stable. The advantages of pooled sera are: (a) its low cost; (b) its lack of errors during reconstitution; and (c) it is the same material as the patient samples. The major disadvantages are potential hepatitis risks and unstable analytes requiring frequent pooling.

Commercially assayed sera can be purchased in a lyophilized form which must be reconstituted with distilled water before use. These sera are stable and have a long shelf life so that
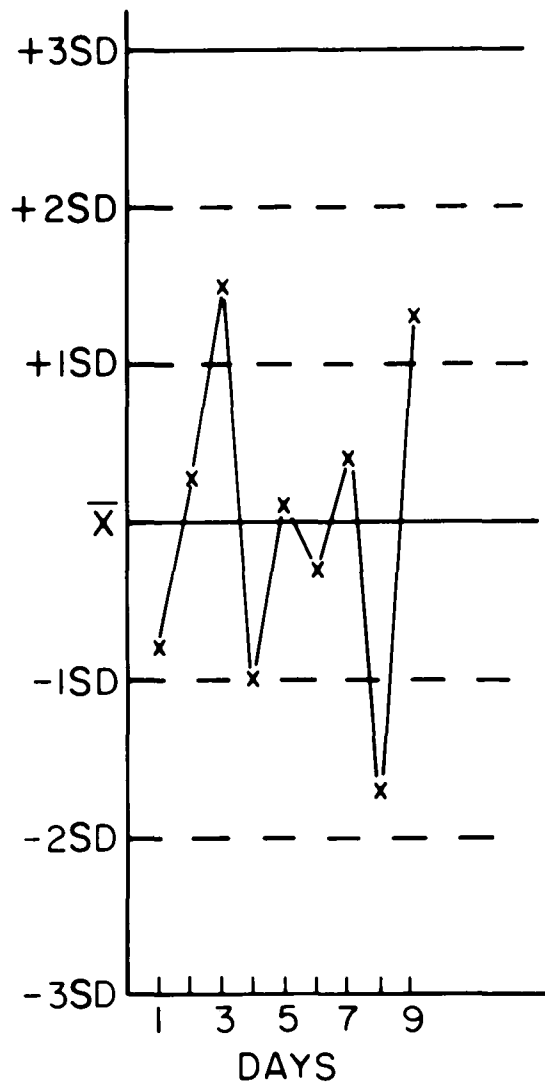

FIG. 2. On a Levey-Jennings Chart the mean and ± 3 s.d. limits are plotted for each control serum. These limits are initially established on at least 20 replicate measurements and periodically updated. Accepting or rejecting runs are often based upon these charts.

they can be used for several assays. Another advantage is that they have a predicted value. The major disadvantage of commercial control sera is that the protein matrix may differ from patient samples causing the controls to behave differently in the assay.

Prepared in-house controls can be made by weighing out a known amount of pure analyte, preparing a solution of proper concentration, and adding it to either serum stripped of the endogenous ligand by charcoal or a serum pool with a low concentration of the ligand. This control allows you to monitor accuracy and precision simultaneously. However, an in-house control is time consuming and expensive to prepare. In addition, some pure analytes may be difficult to obtain.

In selecting the proper control source, the control concentration must be chosen at several levels of the standard curve to cover the range of clinically important concentrations. It is important to monitor the precision of every clinical decision-making range (usually low, normal, and high values) (5).

Once the proper control sera have been prepared, control limits must be established by assaying one aliquot of each control material for at least 20 runs. Calculate the mean, s.d., and CV to determine the allowable limits of variation between runs (interassay precision). The mean and standard deviations are used in the preparation of quality control charts initially described by Shewhart (6) and later introduced into clinical chemistry by Levey and Jennings (7). Control data can then be displayed on control charts each time an assay is run (Fig. 2) so that shifts, drifts, and trends can be easily recognized. Accepting patient results for reporting or rejection of the run

for troubleshooting is often based upon evaluation of these charts.

When results do not fall within $\pm$ 2 s.d. of established ranges, the run is often rejected. However, on occasion, the integrity of the quality control samples may be questionable. Ideally, the QC material is identical to the patient sample in the matrix and analyte to be measured. In practice, however, this is not necessarily the case. Commercial controls are sometimes human or animal serum pools that are spiked and then freeze-dried. Laboratories frequently aliquot the reconstituted controls into convenient portions to be frozen and thawed as needed. The intial process of freeze-drying, reconstitution, followed by another freeze/thaw cycle, renders the QC material less stable. Samples stored as aliquots in small amounts will evaporate (especially in automatic defrost freezers) so that over time they become more concentrated giving an upward trend. This occurrence may lead the technologist to a false impression of reagent deterioration. Because home-made pools of patient samples do not contain stabilizers, they can also show trends or shifts.

CALIBRATION PARAMETERS

A feature of any QC program is maintaining a record of certain standard curve parameters as performance markers. These parameters should be fairly constant from one run to the next and can indicate information about the reagents, environmental conditions, and equipment used in the assay. Control records should include parameters such as total counts, nonspecific binding (NSB), maximum binding, and 50% max-

TABLE 1. Calibration Parameters, Quality Control Effects, and Possible Causes

Parameter	Effect	Causes
Nonspecific binding	Low	Lot changed; improved performance
	High	Change in standard matrix; old tracer; deteriorated tracer; failure in separation system
Maximum binding	High	Antibody titer high; separation failure
	Low	Degradation of tracer; antibody titer low; separation failure
ED50	Change	Antibody concentration change; environmental condition change; separation failure
Control serums	Shifts	Deteriorated standards or tracer; improperly stored controls; instrument changes; reagent lot changes
	Trends	Deteriorated reagents; lot change of standards; improperly stored controls; instrument problems
Patient mean	High/Low	Improper collection; improper storage; incomplete pretreatment; contamination

imum binding (ED50) (8). Table 1 lists each parameter, the QC effect, and the possible cause.

TOTAL COUNTS

Replicate radioactive samples using the same volume as that of other tubes in the assay may cause errors in volume dispensing of the tracer. The mean count rate should remain relatively constant except for radioactive decay and agreement among replicates reflects pipetting precision. Knowing the total counts enables Scatchard analysis of the data and maximum binding calculations.

Nonspecific Binding

The nonspecific binding tubes contain all reagents except the primary binder. The percent bound indicates the degree of binding of any tracer that is not caused by specific ligand-antibody reaction and that may be affected by incubation conditions or separation technique. Labeled materials often have different affinities for various surfaces. Consequently, a change in test tube type can affect the amount of NSB. Ideally, it should not exceed 5%. Increased NSB indicates a decrease in sensitivity and usually causes falsely elevated values.

Maximum Binding

Maximum binding occurs in the absence of an unlabeled ligand and is calculated from replicate tubes containing all reagents except the zero standard that is substituted for the patient sample. These tubes indicate how much tracer is bound by the antisera and may decrease from the ideal 50% for several reasons (i.e., high NSB, low binding affinity or titer of the antiserum, incomplete separation of bound from free ligand, failure to reach equilibrium, or degradation of labeled antigen). Maximum binding may increase if the antibody titer is too high or separation is incomplete.

50% Maximum Binding

The dose corresponding to 50% maximum binding is referred to as the ED50. It is the region of the curve with greatest precision and should approximately lie between one-fourth and one-half of the concentration range. This standard curve parameter is very stable and perturbation can be caused by any factor influencing the reagents (change in antibody concentration), separation system failure, and changes in reaction conditions such as temperature and timing.

EQUIPMENT CONTROL

A comprehensive QC program should cover every aspect of laboratory work. Changes in equipment can produce easily detected changes in test results (9). A gradual slowing of the centrifuge speed or changes in water bath temperature can affect an assay's precision. Daily records of water bath temperatures, refrigerator/freezer temperatures, and counting instrument stability must be maintained. All semiautomatic pipettes (and other tools used for measurement) must be checked for accuracy and precision at the time of purchase and

TABLE 2. RIA Equipment and Problems

Equipment	Problem	Possible Cause of Problem
Gamma counter	Spurious results; poor precision	Contaminated carrier
	Fewer counts causing greater counting errors or longer counting times	Loss in crystal efficiency; Shift in radioisotope peak
Pipettes	Change in maximum binding and total counts; poor duplication	Poor precision
	Imprecision; spurious results	Bad tips
Water bath	Decreased maximum binding; change in equilibrium point	Increase in temperature
	Increased maximum binding; change in equilibrium point	Decrease in temperature
Centrifuge	Poor pellet formation; imprecision	Speed variation; timer wrong
Mixer	Incomplete suspensions; imprecision	Inefficient
Test tubes	Hydrostatics change affection decanting	Change in type
	Poor mixing or evaporation	Wrong size

periodically thereafter. Tachometer checks of centrifuges must be performed for correct revolutions per minute, and centrifuge timer and temperature calibration should be done quarterly. Even a change in the type of test tube can affect the assay results. Furthermore, decanting efficiency may be altered due to different hydrostatics of the interaction between the reagents and the tube wall. Table 2 lists some RIA equipment and problems frequently encountered.

SAMPLE CONTROL

The condition of the patient samples to be assayed may affect test results. Generally, the daily patient mean is constant (10) and can reflect how the specimens have been handled. Shifts in the mean indicate a systematic problem involving factors other than the assay. A run might show the standards and controls as regulators, but the patient mean may be high (or low). In this case, suspect either an error in collection (i.e., the wrong tube type or the wrong time of day) or an error in specimen handling. Analytes that are unstable can cause lower daily means if stored improperly. Assays, which require a sample denaturation step to release analyte from endogenous binders such as boiling, may have boiled too long or not long enough, producing spurious results. Contamination of the samples or diluents with bacteria can shift the mean of the results.

Individual sample results can be affected by endogenous contaminants such as hemoglobulin, bilirubin, or lipoproteins. The condition of the sample should be noted upon its arrival in the laboratory. Each assay performed in the laboratory should be evaluated for the effects of such interferences, so that results can either be reported or another sample can be requested.

When an individual sample exhibits an unpredictable value, a parallelism study can be quickly performed to reveal the presence of an interference in the patient's sample (11). Cross-reacting compounds and matrix effects such as pH, salts, and endogenous binders can interfere with antibody binding, thereby reducing the apparent ligand concentration. These interferences can also affect the separation of the bound and free components. Dilutions of the sample are made with the zero standard in a 1:2, 1:5, 1:10, and 1:20 concentration. The assay is peformed on each dilution and undiluted sample in the same run. The amount expected in each dilution is calculated from the value obtained on the undiluted sample and the dilution factor. A plot is then made of Expected Results as compared to Observed Results (Fig. 3). A visual inspection will quickly demonstrate nonparallelism.

SYSTEMATIC APPROACH TO TROUBLESHOOTING

The first step in the process of troubleshooting an assay is the identification of a problem. After each run is performed, the calibration curve, the quality control samples, and the patient samples must all be inspected. Compare the calibration curve parameters such as total counts, maximum binding, NSB, and ED50 to the record you have been keeping. Quality

control samples should all fall within ± 2 s.d. of the established means. If one control falls between 2 and 3 s.d. for the first time, the supervisor should be notified and the results released. If two controls fall between 2 and 3 s.d., if one control is greater than 2 s.d. on two successive runs, or if one control is outside ± 3 s.d., results should be held and troubleshooting action started (12). Shifts, drifts, and trends must also be noted. Frequently, shifts can occur only in certain sections of the concentration range. This situation can happen if one standard is improperly reconstituted or contaminated. The patient sample mean should also be noted and compared to the track record.

Once a problem has been identified, gather together all involved personnel (i.e., technologist performing the assay, the lead technologist, and the supervisor). A few minutes spent discussing the problem can save hours of frustration later. Ordinary mistakes can usually be found at this point by checking the obvious, such as omitted reagents, use of incorrect pipet, the wrong reaction time or temperature, or incorrectly reconstituted standards or controls.

When these criteria cannot explain why control values are beyond limits or the standard curve parameters are off, other possibilities must be listed and ruled out until only one remains. These problem variables can be grouped into three areas: assay reagents, environmental variable, and equipment.

Any reagent used in the assay must be considered, and the tracer and antibody should be checked for changes in lot number. A low maximum binding with a high NSB may indicate a damaged tracer. A change in the ED50 may indicate a change in antibody concentration. Consultations with the manufacturer's technical support staff may help to determine if the antibody dilution was changed or if other laboratories have reported similar problems. Standards and quality control samples should be inspected for physical signs of deteriora-
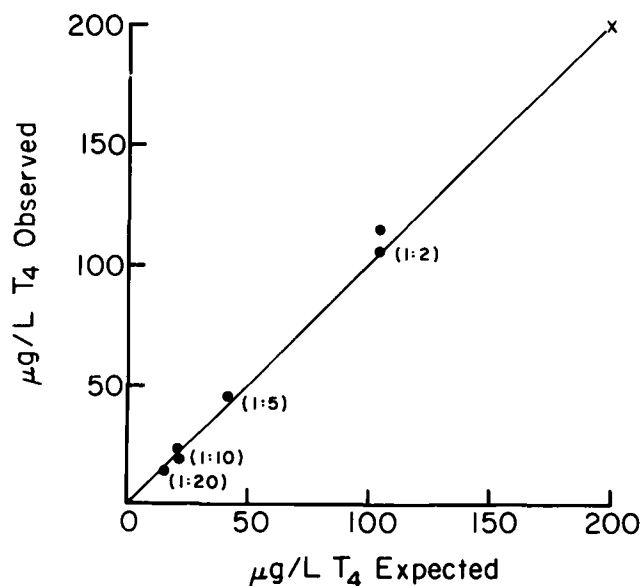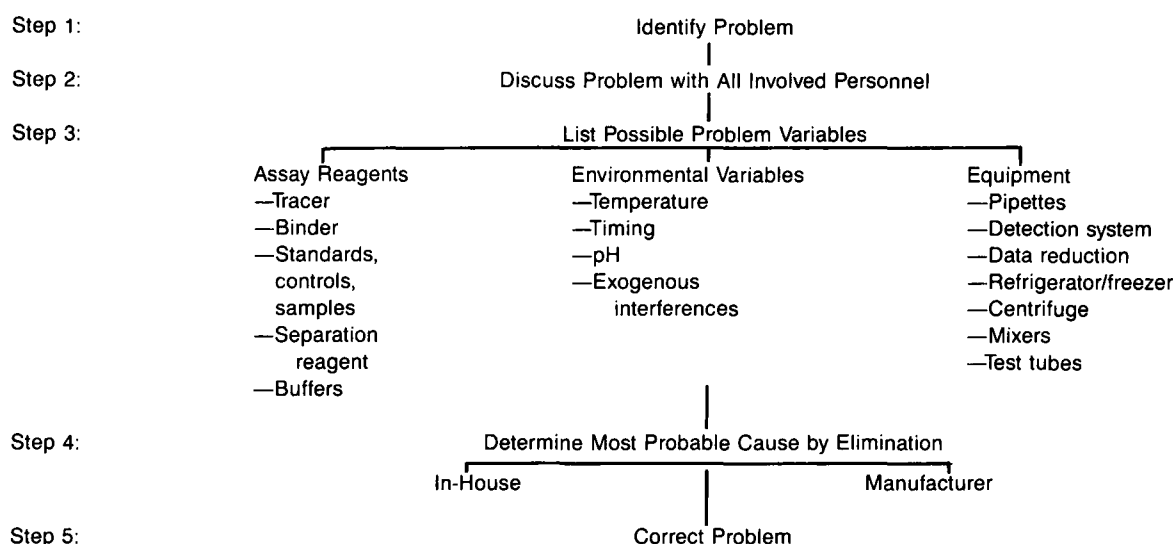

FIG. 3. Parallelism plot of Expected as compared to Observed results. Ideally, a straight line should be drawn with a slope of one, an intercept of zero, and a correlation coefficient of 1.0.

168

TABLE 3. Systematic Approach to Troubleshooting

Step 1:	Identify Problem	

Step 2:   Discuss Problem with All Involved Personnel

Step 3:   List Possible Problem Variables

Assay Reagents	Environmental Variables	Equipment
—Tracer	—Temperature	—Pipettes
—Binder	—Timing	—Detection system
—Standards,	—pH	—Data reduction
controls,	—Exogenous	—Refrigerator/freezer
samples	interferences	—Centrifuge
—Separation		—Mixers
reagent		—Test tubes
—Buffers		

Step 4:   Determine Most Probable Cause by Elimination

In-House   Manufacturer

Step 5:   Correct Problem

tion. If a computer data reduction program is used, standard concentrations should be checked for calibration value changes. Separation reagents should be inspected for lot changes, which can trigger problems in the high concentration range caused by instability or changes in NSB. Assay buffers should also be physically examined for contamination and pH changes.

All assays are sensitive to environmental variables such as pH, time, and temperature. Nonequilibrium assays and those which work by sequential addition are extremely sensitive to time, whereas equilibrium assays are less so. When large assays are performed, control sera must be placed intermittently throughout the assay to ensure uniform timing of Tube 1 and Tube 300. In addition, reaction temperatures must be carefully maintained. Perhaps the most difficult reaction to control is that which occurs at room temperature. Laboratory temperatures can change significantly from one day to the next, and incubation of test tubes on a sunny window sill or under an air conditioning vent must be carefully avoided.

Exogenous interferences are more likely to occur in individual serum samples. Sample inspection and a parallelism study will help rule out endogenous binders. Radioactive contamination from in vivo nuclear medicine studies may be suspected and ruled out by simply counting an aliquot of the sample.

Quality control records of equipment used in assay performance should be checked. Furthermore, pipets should be rechecked for precision and accuracy. Counting instruments should be checked for counting efficiency, contaminated carriers, or contaminated crystal.

By troubleshooting each of the possible causes, the most probable cause can be determined by the process of elimination. The cause may be an in-house problem involving the equipment or the environment. If that is the case, corrective

action should then be taken before repeating the assay. If the problem involves the kit reagents, the manufacturer should be contacted and is then responsible for correcting the problem. If the manufacturer is unable or unwilling to correct the problem, changing kit manufacturers may be a possible solution. Table 3 shows a diagram outlining an easy and systematic approach to answer the question *why* when your assay fails to meet your preset quality control criteria.

SUMMARY

Reliable patient results are ensured daily with a good quality control program that indicates correct and incorrect operations and allows the technologist to determine error sources that produce undesirable test results. In addition, monitoring standard curve parameters, quality control sera, and patient sample means, will increase confidence in the quality of the results. Furthermore, a troubleshooting process that incorporates a systematic approach while utilizing each facet of the quality control program can direct the technologist to the most probable cause of a problem by the process of elimination.

REFERENCES

1. Ekins R. The "precision profile": Its use in RIA assessment and design. *Ligand Quarterly* 1981;2:33–44.

2. Feldkamp C. Immunoassay kit evaluation and selection. *J Nucl Med Technol* 1985;13:88–94.

3. Westgard JO, Barry PL, Hunt MR, et al. A multi-role Shewhart chart for quality control in clinical chemistry. *Clin Chem* 1981;27:493–501.

4. Jeffcoate SL. Quality control and standardization of hormone assays in human reproduction. *J Clin Immunoassay* 1983;1:68–75.

5. Monroe L, Thompson W, Kelley E, et al. Quality control in the in vitro nuclear medicine laboratory. Presented at the 26th Annual Meeting of The Society of Nuclear Medicine, Anaheim, CA, June 27, 1979.

6. Shewhart WA. *Economic Control of Quality in Manufactured Products.* New York: Van Nostrand, 1931.

7. Levey S, Jennings ER. The use of control charts in the clinical laboratory. *Am J Clin Pathol* 1950;20:1059–66.

8. Rodbard D, Rayford PL, Cooper JA. Statistical quality control of radio-immunoassays. *J Clin Endocrin and Metab* 1968;28:1412–18.

9. Perlstein M. Sensible QC and troubleshooting for the RIA lab. Corning Medical Technical Bulletin, August 2, 1982.

10. Hoffman RG, Waid ME. The "average of normals" method of quality control. *Am J Clin Pathol* 1965;43:134–41.

11. Perlstein MT, Chan DW, Bill MJ. Parallelism: A useful tool for troubleshooting. *Ligand Quarterly* 1980;1:34–36.

12. Perlstein MT. A problem solving primer: A practical approach to QC and troubleshooting in immunoassays. Corning Medical.