Quality control for qualitative assays: quantitative QC procedure designed to assure analytical quality required for an ELISA of hepatitis B surface antigen

GEORGE A. GREEN, IV, 1* R. NEILL CAREY, 2 JAMES O. WESTGARD, 3 TAMi CARTEN, 1 LEEANN SHABLESKY, 1 DANIEL ACHORD, 1 EILEEN PAGE, 1 and ANH VAN LE 1

An assay for hepatitis B surface antigen (HBsAg) should reliably detect 0.2 µg/L, the lowest reported concentration in an asymptomatic blood donor. The difference between this concentration and the assay cutoff defines the analytical quality requirement in a total error format. The design of a statistical QC procedure is critically dependent on the precision of the assay. The precision of a developmental ELISA of HBsAg under study ranged from 17.5% to 9.6% for controls containing 0.07 to 1.50 µg/L, respectively. Use of one positive control with the 13s QC rule provided an 85% chance of detecting a critical loss of assay sensitivity; use of two positive controls increased the chance of detecting critical loss of assay sensitivity to nearly 100%. These rules are based on the precision of this developmental assay, and must be developed individually for other assays. The development of the proposed QC procedures illustrates how quantitative QC can be provided for qualitative assays.

Conventional QC procedures for ELISAs of hepatitis B surface antigen (HBsAg) 4 rely on testing strongly positive controls included in the reagent kit. Often the acceptability ranges for the resulting absorbance values are broad, to be more widely applicable. QC procedures operating with these controls and wide ranges may not be able to detect a clinically significant loss of assay sensitivity. The Clinical Laboratory Improvement Act of 1988 requires that laboratories use positive controls separate from those used to calculate the cutoff [1], and some states require use of a positive control in addition to the reagent kit manufacturer’s test kit controls [2]. Without guidance on how to implement these new controls, there is no guarantee that QC of the assay will be improved by these regulations. Here, using an ELISA of HBsAg capable of detecting <0.1 µg/L, we considered how a user laboratory could design an improved QC procedure and what level of performance could be expected from implementing it.

The first step in the design of a QC procedure is to define the quality requirements of the assay [3]. For HBsAg, the clinical decision value is at the cutoff concentration; however, the lowest reported HBsAg concentration in an asymptomatic blood donor (0.2 µg/L) is far greater than the cutoff in this assay [4]. Virtually all subjects who are infected with hepatitis B will test very strongly positive for HBsAg soon after infection, at concentrations tens to hundreds of times the cutoff. Thus, the difference between 0.2 µg/L and the cutoff defines the analytical quality requirement in a total error format. Any run in which 0.2 µg/L HBsAg fails to produce a signal above cutoff should be rejected.

The next step in the design of a QC procedure is to evaluate the precision of the assay. We used an analysis of variance experiment to study assay precision at concentrations from 0.07 to 1.50 µg/L.

Finally, we selected appropriate control rules and numbers of control measurements on the basis of the information about the quality requirement, the observed method performance characteristics, and the performance characteristics of the candidate QC procedures, which were determined by computer simulation studies.
Materials and Methods

ELISA for HBsAg and positive controls. We used the Ortho Antibody to HBsAg ELISA Test System 3 (Ortho Diagnostics Systems). The assay was automated with the Ortho Summit™ sample-handling system, operated with Ortho ELISA ver. 5.2C™ software. Plates were incubated in Ortho Model 120 forced-air incubators, washed with an Ortho Auto Wash 96™ plate-washer, and read with an Ortho AutoReader II™. We made positive controls with variations by computer simulation with Minitab as previously stated. We calculate the total standard deviation (s) for each control, used Minitab (State College, PA) statistical software to determine from performance data of the assay. We derived from the mean absorbance of three negative controls, the cutoff absorbance was determined. As with most qualitative tests, the ratio of sample absorbance to cutoff absorbance (S/C) has the effect of compensating for between-run variations. For moderately positive controls, however, the absorbance tends to compensate for between-run variations. For moderately positive controls, however, CV is nearly 1% greater when results were expressed as S/C (10.4% at 1.50 μg/L HBsAg). CV, calculated from S/C data is lower for low-positive controls because absorbance values of similar magnitude are being ratioed, and dividing by the cutoff absorbance tends to compensate for between-run variations. For moderately positive controls, however, the sample absorbance is >10-fold the cutoff absorbance and has a lower CV, and between-run variation is a smaller proportion of the total variation. Dividing the relatively more precise moderately positive control absorbance by the relatively less precise cutoff absorbance leads to the increase in CV.

Power functions were then entered into the QC Validator program (Westgard QC, Ogunquit, ME) by using the program’s utility for editing the candidate QC file. Critical-error graphs and charts of operating specifications (OPS specs charts) were then prepared with the QC Validator program. For candidate QC procedures, we selected rules that are relatively easy to implement and have reasonably low false-rejection rates, e.g., $2\times 0.01, 3\times 0.01, 2\times 0.002, 3\times 0.002, X_{0.01}/R_{0.01}$, and $X_{0.002}/R_{0.002}$ rules. For calculation of control limits for these rules, see Westgard et al. [7].

Results

Mean absorbance values, within-plate CVs ($CV_w$), total CVs ($CV_t$), and ratios of between-run to within-run CVs are given in Table 2 for each control. When control data are expressed in terms of absorbance, total CVs decrease with increasing concentration, from 17.2% at 0.07 μg/L HBsAg to 9.6% at 1.50 μg/L. Expressing data in terms of S/C has the effect of reducing the between-run variations for lower-concentration controls and reduces $CV_t$ (13.3% at 0.07 μg/L HBsAg). For the two moderately positive controls, however, $CV_t$ is nearly 1% greater when results were expressed as S/C (10.4% at 1.50 μg/L HBsAg). $CV_t$ calculated from S/C data is lower for low-positive controls because absorbance values of similar magnitude are being ratioed, and dividing by the cutoff absorbance tends to compensate for between-run variations. For moderately positive controls, however, the sample absorbance is >10-fold the cutoff absorbance and has a lower CV, and between-run variation is a smaller proportion of the total variation. Dividing the relatively more precise moderately positive control absorbance by the relatively less precise cutoff absorbance leads to the increase in CV.

Although we did not test a control exactly at the medical decision concentration (0.2 μg/L), the $CV_t$ of the closest-concentration control, 0.1 μg/L, is a suitable estimate of the $CV_t$ at 0.2 μg/L for the purpose of designing a QC procedure. The $CV_t$ at 0.2 μg/L expressed in absorbance should be slightly less than that at 0.1 μg/L, whereas the $CV_t$ at 0.2 μg/L expressed in S/C should be nearly equal to that at 0.1 μg/L.

Because the response of the assay is linear, we calculated the expected absorbance value for 0.2 μg/L HBsAg from the absorbance of the 0.1 μg/L control and obtained 0.118. The mean cutoff absorbance was 0.034. Thus, the total allowable error is a loss of assay sensitivity from 0.2 μg/L HBsAg to 0.118 μg/L (0.118 A) down to cutoff (0.034 A), or a 71.2% loss of sensitivity.

We first assessed the QC design for data expressed in terms of absorbance. The critical systematic error, $△S_{crit}$, was calculated [8] from the allowable error and the $CV_t$ of the 0.1 μg/L control to be a shift equivalent to 3.26 s. Power function curves for systematic error (loss of sensitivity) are shown for several candidate control rules, with use of 1 or 2 positive controls per run, in Fig. 1. All rules with $P_{ed}$ near 0.90 had $P_{fr}$ of 0.02 or greater. For a

---

**Table 1. Plate (96-well) layout for precision evaluation.**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>BL</td>
<td>0.07°</td>
<td>1.50</td>
<td>1.00</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>B</td>
<td>NS</td>
<td>0.07°</td>
<td>1.50</td>
<td>1.00</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>C</td>
<td>NS</td>
<td>0.10°</td>
<td>1.50</td>
<td>1.00</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>D</td>
<td>NS</td>
<td>0.10°</td>
<td>1.50</td>
<td>1.00</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>E</td>
<td>1.50°</td>
<td>1.50</td>
<td>1.00</td>
<td>0.07</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>1.50°</td>
<td>1.50</td>
<td>1.00</td>
<td>0.07</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>1.00°</td>
<td>1.50</td>
<td>1.00</td>
<td>0.07</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>1.00°</td>
<td>1.50</td>
<td>1.00</td>
<td>0.07</td>
<td>0.10</td>
<td></td>
</tr>
</tbody>
</table>

*a* Last four wells in column 6, and all wells in columns 7–12, contained negative specimens.

*b* Control used for plate validation.

BL, blank; NS, negative standard.
reasonable tradeoff of QC rule sensitivity to error and low false rejections, the best choice is probably the $1_{3s}$ rule with 2 positive controls, which has $P_{rd}$ of 0.81 and $P_{fr}$ near 0.00.

Mean and range rules demonstrated high sensitivity to shifts; however, they are not shown on Fig. 1 because of their high $P_{fr}$ of 0.04 and 0.02, respectively, for the $X_{0.01}/R_{0.01}$ and $X_{0.002}/R_{0.002}$ rules. Between-run variations increase the $P_{fr}$ for control rules that rely on the mean, and generally reduce sensitivity to systematic error.

Expression of control data in terms of S/C reduced the $CV_w$, increasing the critical systematic error from 3.26 s to 4.09 s. In Fig. 1, a critical systematic error of 4.09 is off-scale to the right; all rules have slightly higher $P_{fr}$ than shown. Even the $1_{3s}$ rule has very high sensitivity to the critical systematic error when 2 positive controls are run. Acceptable sensitivity to systematic error is now provided by the $1_{3s}$ rule with only 1 positive control per run.

**Discussion**

We used four controls to evaluate the precision performance of the assay. The 0.07 and 0.10 μg/L HBsAg controls demonstrated the precision between the assay cutoff and the 0.20 μg/L decision concentration. Concentrations of 1.00 and 1.50 μg/L were typical of the moderately positive controls included in commercial HBsAg reagent kits. To obtain optimal performance from QC procedures when low-positive controls are used, one should express control results in S/C. For moderately positive controls, control results can be expressed in either S/C or simple absorbance because the between-run variations have proportionately less impact on simple absorbance as the concentration of the control increases. Moderately positive controls may be preferable in practice because QC ranges calculated from absorbance are simpler to implement than those calculated from S/C.

When results were expressed in terms of S/C, high probability to detect loss of assay sensitivity was demonstrated by both the $1_{3s}$ rule with 2 positive controls per run and the $1_{3s}$ rule with 1 positive control per run. Both rules have $P_{fr}$ of <0.01 and $P_{ed}$ >0.85, even when simulations take into account between-run variations. With 2 positive controls per run, the $1_{3s}$ rule is virtually certain to detect critical loss of assay sensitivity; $P_{fr}$ is <0.01. Because the 0.1 μg/L control tested positive on all replicates on all plates, the frequency of unacceptable runs is expected to be very low when this ELISA is in routine use. Thus, QC procedures with 1 positive control and having $P_{ed}$ slightly <0.90 would be satisfactory.

The HBsAg ELISA is a single-analyte assay. Assays for other infectious disease markers, such as human immunodeficiency virus, human T-lymphocyte virus, and hepatitis C virus, detect multiple analytes and thus may not give a direct linear response of controls to assay sensitivity. Further study is necessary to determine if the methods used here will apply to other ELISAs.

Because this experiment was designed to study the performance characteristics of positive controls, we did not run negative controls other than the negative materials used to calculate the cutoff absorbance. In ELISAs for HBsAg, absorbance values for negative controls are very close to 0 (typically ranging from <0 to 0.005 A). Because reactive specimens are repeated in duplicate in a second run, a negative sample must test falsely positive 3 times in 2 runs to be reported as reactive for HBsAg. Thus, negative controls function mostly to detect gross contamination and the presence of splashing or poor washing during the assay.

The assay protocol for Ortho Antibody to HBsAg ELISA Test System 3 specifies running 3 negative calibrators and uses a test of their range, as well as a test of their mean absorbance, to judge the acceptability of their absorbance values before their mean absorbance is used to detect loss of assay sensitivity. The power function curves for systematic error of HBsAg ELISA are shown in Fig. 1.

**Table 2. Summary of precision data.**

<table>
<thead>
<tr>
<th>Control concn., μg/L</th>
<th>Absorbance</th>
<th>S/C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>CVw</td>
</tr>
<tr>
<td>0.07</td>
<td>0.041</td>
<td>10.1</td>
</tr>
<tr>
<td>0.10</td>
<td>0.059</td>
<td>10.2</td>
</tr>
<tr>
<td>1.00</td>
<td>0.596</td>
<td>8.7</td>
</tr>
<tr>
<td>1.50</td>
<td>0.874</td>
<td>8.4</td>
</tr>
</tbody>
</table>

CVw, CV within-run; CVt, CV total; CVb, CV between-runs.
calculate the cutoff absorbance. In this case, the negative calibrators provide the same information as negative controls. In our experience, nonrepeatable false positives usually occur from assignable causes, such as splashing from an adjacent well. Reference sample QC has limited ability to detect random conditions that favor these types of false positives. We did not study the ability of reference sample QC to detect an systematic increase in sensitivity (e.g., causing negative samples to test positive). Because the mean of HBsAg-negative results is >6 s from the cutoff, such an increase in sensitivity will only rarely cause a false-positive result. This may not hold true for other ELISAs for hepatitis markers, in which the absorbance of negative controls has considerable between-run variance.

These studies demonstrate that QC procedures that guarantee a stated amount of sensitivity can be developed for qualitative HBsAg ELISAs. These rules can be selected to give a high probability of error detection while maintaining low probabilities of false rejection. In practice, it will be necessary for manufacturers to continue to provide the general QC limits that would apply to all laboratories in initially establishing an assay. User laboratories can improve the performance of the assay by determining their own quality requirements, evaluating assay precision in their own environments, and developing custom QC rules, as is the practice with most clinical chemistry assays.

We appreciate the technical assistance provided by Marla Fosdick, Michelle Anson, and Susan Merkel of the Blood Bank of Delaware. R.N.C. is a consultant to Ortho Diagnostic Systems, and J.O.W. is a principal in Westgard QC.

References