CLIN. CHEM. 39/3, 440–447 (1993)

New Insight into the Comparative Power of Quality-Control Rules That Use Control Observations within a Single Analytical Run

Curtis A. Parvin

The error detection characteristics of quality-control (QC) rules that use control observations within a single analytical run are investigated. Unlike the evaluation of QC rules that span multiple analytical runs, most of the fundamental results regarding the performance of QC rules applied within a single analytical run can be obtained from statistical theory, without the need for simulation studies. The case of two control observations per run is investigated for ease of graphical display, but the conclusions can be extended to more than two control observations per run. Results are summarized in a graphical format that offers many interesting insights into the relations among the various QC rules. The graphs provide heuristic support to the theoretical conclusions that no QC rule is best under all error conditions, but the multirule that combines the mean rule and a within-run standard deviation rule offers an attractive compromise.

Indexing Terms: *statistics · computer modeling · error detection*

Evaluation of the ability of a quality-control (QC) rule to detect out-of-control states that persist until they are discovered and corrected was the topic of a recent series of papers (*1–3*). These papers showed that for the evaluation and comparison of QC rules that span more than a single analytical run, the simulation approach that has been used routinely in the laboratory medicine literature is not appropriate (*4–6*). This approach estimates the irrelevant probability of rejecting an analytical run independently of whether previous runs are accepted or rejected. The concepts of probability of error detection (p_{ed}) and probability of false rejection (p_{fr}) must be replaced by concepts such as the average run length (ARL) to rejection, the average conditional probability of rejection, or the cumulative probability of rejection at the *i*th run with error (*1*). Estimates of these alternative concepts can be obtained from simulation approaches that allow each cycle of the simulation to test successive analytical runs until a run is rejected. Simulation is an essential tool in this situation because of the complexity of the QC rules and the lack of independence of these rules from run to run, making attempts to derive statistical theory very difficult.

The purpose of this paper is to investigate the error detection ability of rules that only use control observations within the current run. Many of the QC rules that have been proposed and evaluated, and that are used routinely in the clinical laboratory, utilize control observations in the current analytical run only. The concepts of probability of error detection and probability of false rejection are meaningful and useful in this situation; either the traditional or alternative simulation approach can be used to evaluate error detection ability. However, in contrast to the situation where QC rules are applied across analytical runs, rules that apply only to the current run make it possible to draw many important fundamental conclusions by using statistical theory, without the need to simulate. A few papers have pointed to some of this theory within the context of QC in clinical chemistry (*7, 8*). In this paper, the theory is extended and conveyed within a graphical context that offers intuitive support to the findings.

Six different QC rules are compared. By Westgard's notation (*9*) they are the 1_{3s} rule, 2_{2s} rule, R_{4s} rule, $\bar{X}$ rule, $1_{3s}/2_{2s}/R_{4s}$ rule, and $\bar{X}/R_{4s}$ rule. The case with two control observations per run will be used throughout, thereby allowing convenient graphical presentation of the results. All evaluations are performed without simulation. The results are summarized by displaying two-dimensional graphs showing contour plots for the bivariate distribution of the control observations when the process is in control or, when it is out of control, superimposed on a two-dimensional representation of the rejection region for a particular rule. Displayed in this way, the graphs provide interesting insights into the relations among the various QC rules and their comparative ability for error detection. The results generalize in a straightforward way to more than two control observations per run.

Methods

I assume there are two control observations per run and that the control observations follow the model

$$C_{ij} = \mu_j + \sigma_{b_j} E_i + \sigma_{w_j} e_{ij}, \quad i = 1, 2, 3, \ldots, j = 1, 2$$

where C_{ij} denotes the measured value of the *j*th control observation in the *i*th analytical run, E_i represents random between-run error, and e_{ij} represents random within-run error for the *j*th control observation in the *i*th analytical run. During stable operation, the distributions of E_i and e_{ij} are assumed to be independent and Gaussian with means = 0 and variances = 1. For a specific analytical run (say, run 1) the between-run error (E_1) is constant and the two control measurements within the run are independently distributed with

Division of Laboratory Medicine, Departments of Pathology and Medicine, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, MO 63110.

[1] Nonstandard abbreviations: QC, quality control; SE, systematic error.

Received June 9, 1992; accepted October 13, 1992.

means $\mu_1 + E_1$ and $\mu_2 + E_1$ and variances $\sigma^2_{w_1}$ and $\sigma^2_{w_2}$. However, for an arbitrary analytical run (say, run i) E_i is random and the total imprecision of the control measurements must reflect the added imprecision due to the random between-run error. Additionally, the control measurements are not independently distributed, but are correlated because of the common component of between-run error that they share. Therefore, in the general case, C_{i1} and C_{i2} follow a bivariate Gaussian distribution with means μ_1 and μ_2, variances $\sigma^2_{t_1} = \sigma^2_{b_1} + \sigma^2_{w_1}$ and $\sigma^2_{t_2}$, and correlation (ρ) between C_{i1} and C_{i2} of $(\sigma_{b_1}\sigma_{b_2})/(\sigma_{t_1}\sigma_{t_2})$. The control rules evaluated in this paper use control observations within a single analytical run; therefore, to streamline notation, the run subscript, i, will be dropped.

This model is a more general formulation of the model I described previously (*3*). If the two control observations in a run are on the same control material, then $\mu_1 = \mu_2$, $\sigma_{b_1} = \sigma_{b_2}$ and $\sigma_{w_1} = \sigma_{w_2}$. If the two control observations are on different control materials, then $\mu_1 \neq \mu_2$, and the magnitudes of between-run and within-run imprecision for the two control materials may or may not be equal.

The ratio of between-run to within-run standard deviation for the jth control observation during stable operation will be denoted ϕ_{s_j}. It is easy to verify that the relation of ρ to ϕ_{s_1} and ϕ_{s_2} is

$$\rho = \sqrt{\left(\frac{\phi^2_{s_1}}{1+\phi^2_{s_1}}\right)\left(\frac{\phi^2_{s_2}}{1+\phi^2_{s_2}}\right)}$$

Two situations are evaluated; the case where there is no between-run error ($\phi_{s_1} = \phi_{s_2} = 0$), which implies $\rho = 0$, and the case where the magnitudes of between-run and within-run imprecision are equal ($\phi_{s_1} = \phi_{s_2} = 1$), which implies $\rho = 0.5$. For simplicity, these cases will be referred to as $\phi_s = 0$ and $\phi_s = 1$. Three different error situations are evaluated. In the first case a systematic error (SE) in the QC mean is evaluated. I assume a shift in each QC mean (μ_j) by an amount SE = $2.0\sigma_{t_j}$. Next, I assume that total analytical imprecision increases from its stable value, σ_{t_j}, to $RE_b = 1.5\sigma_{t_j}$, because of an increase in the between-run component of imprecision from its stable value, σ_{b_j}, to an out-of-control value σ_{B_j}. Finally, I assume that total analytical imprecision increases from its stable value, σ_{t_j}, to $RE_w = 1.5\sigma_{t_j}$, because of an increase in the between-run component of imprecision from its stable value, σ_{w_j}, to an out-of-control value, σ_{W_j}. The formulas used to determine the magnitude of σ_{B_j} or σ_{W_j} that increase total imprecision by a factor of 1.5 were described previously (*3*).

All analyses are performed on the transformed variables $z_j = (C_j - \mu_j)/\sigma_{t_j}$, which have means = 0, variances = 1, and the same correlation between z_1 and z_2 as given earlier for C_1 and C_2. The probability of rejection for a rule is obtained by numerically integrating the volume under the appropriate bivariate Gaussian probability density function over the rejection region defined by the rule. The Mathematica software package was used for the numerical integration and for creation of the figures (*10*).

Results

Figure 1 portrays the bivariate Gaussian distributions describing the probability density for z_1 and z_2 when there is no between-run error ($\phi_s = 0.0$, top panel) and when between-run imprecision equals within-run imprecision for each control observation ($\phi_s = 1.0$, bottom panel). The probability density defines a surface in three-dimensional space. For every pair of (z_1, z_2) values there is a density $f(z_1, z_2)$ represented by the height of the surface at that point. Probability corresponds to volume under the surface. Circles ($\rho = 0$) or ellipses ($\rho \neq 0$) define the contours of constant density (elevation) on the surfaces. The following figures represent these bivariate distributions by a single contour

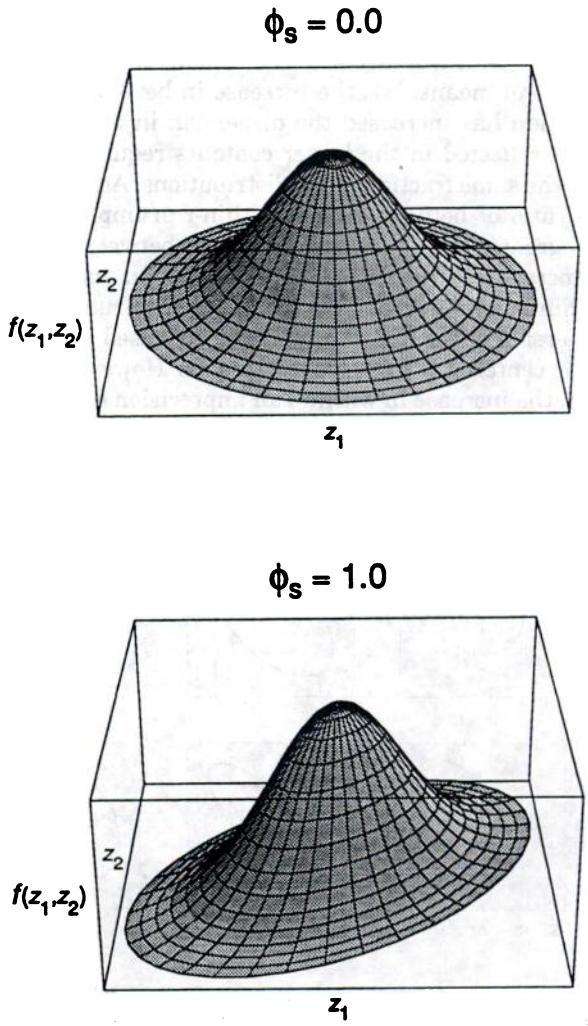

Fig. 1. Surface plots describing the bivariate normal distribution for two control observations within an analytical run

Surface contours trace a constant value for the density function $f(z_1, z_2)$. When no component of between-run imprecision exists ($\phi_s = 0$), z_1 and z_2 are uncorrelated and the contours of the distribution are circles (*top panel*). When between-run imprecision exists ($\phi_s > 0$), z_1 and z_2 are correlated and the contours are ellipses (*bottom panel*)

defined so that the volume contained within it equals a specified probability.

Figure 2 portrays the error detection ability of the 1_{3s} rule. The 1_{3s} rule rejects the run if either one or both of the control observations is >3 total analytical standard deviations away from its mean value. The shaded area in each panel defines the rejection region for the rule. The rejection region extends to $\pm\infty$ along each axis, but, for display purposes, the axes have been restricted to ±6. The dashed contours define the bivariate distribution for z_1 and z_2 during in-control operation. The solid contours define the bivariate distribution under different error conditions. The top three panels are cases where $\phi_s = 0.0$ and the bottom three panels are cases where $\phi_s = 1.0$. The volume contained within each contour equals $1 - p_{fr}$ for the rule when $\phi_s = 0$.

The left panels depict a constant shift in the mean. The out-of-control distributions (solid contours) have the same shape as the in-control distributions (dashed contours), but they are shifted so that they are centered at the out-of-control values for the QC means. The middle panels depict an increase in between-run imprecision. The out-of-control distributions remain centered at the in-control means, but the increase in between-run imprecision has increased the dispersion in the distributions, reflected in the larger contours required to contain the same fraction of the distributions. Additionally, the ratio of between-run to within-run imprecision is now greater than ϕ_s, so the correlation between z_1 and z_2 is increased. The right panels depict an increase in within-run imprecision. As in the between-run case, the dispersion of the distributions has increased while still being centered at the in-control means. However, in this case, the increase in within-run imprecision reduces the ratio of between-run to within-run imprecision below ϕ_s (except when $\phi_s = 0$) and the correlation between z_1 and z_2 is decreased.

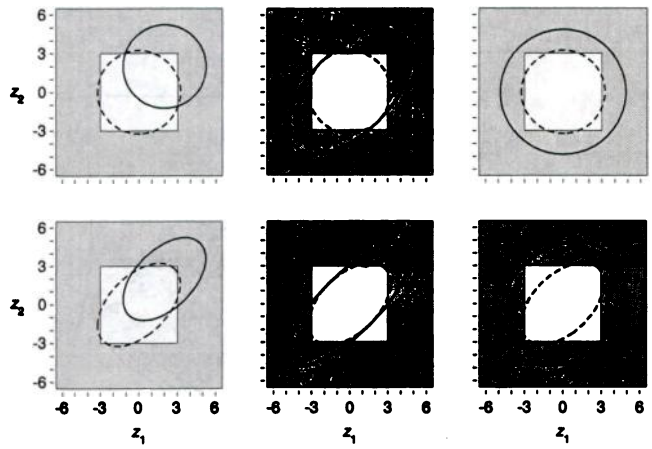

Fig. 2. The error detection characteristics of the 1_{3s} rule

The *shaded area* (extending to $\pm\infty$) denotes the rejection region for the rule. In the *top three panels* no between-run imprecision exists during stable operation. In the *bottom three panels* equal components of between-run and within-run imprecision exist during stable operation. The *dashed contours* represent the bivariate probability distribution of the control observations during in-control operation. The *solid contours* represent the bivariate distribution during three different out-of-control states. *Left panels* represent a shift in the QC means; *middle panels* represent an increase in between-run imprecision; and *right panels* represent an increase in within-run imprecision

The probability that the 1_{3s} rule rejects a run is calculated by integrating the volume under the bivariate probability density describing the current state of the assay over the area defined by the rejection region for the rule. Conversely, the probability that the 1_{3s} rule accepts a run is equal to the volume under the appropriate density function contained within the unshaded region. For instance, the volume contained within the unshaded square region under the bivariate density function represented by the solid circle in the upper left panel of Figure 2 equals 0.7079. This is the probability of accepting the 1_{3s} rule when a shift in the QC means of 2 total analytical standard deviations has occurred. Table 1 lists the rejection probabilities obtained by numerical integration for the in-control and out-of-control states for the 1_{3s} rule depicted in Figure 2 and for the other rules for which figures follow.

Figure 3 depicts the 2_{2s} rule, which rejects the run if both control observations are more than 2 total analytical standard deviations above their means or both are more than 2 total analytical standard deviations below their means. The contours in Figure 3 are larger than those in Figure 2 because, when $\phi_s = 0.0$, the false-rejection rate for the 2_{2s} rule ($p_{fr} = 0.0010$) is smaller than the false-rejection rate for the 1_{3s} rule ($p_{fr} = 0.0054$). Consequently, the contours for the 2_{2s} rule are larger, to contain the additional acceptance probability.

Figure 4 represents the R_{4s} rule, which rejects if the absolute difference between $(C_1 - \mu_1)/\sigma_{w_1}$ and $(C_2 - \mu_2)/\sigma_{w_2}$ is >4. Note that this is the original range rule, which is based on the difference between the two control observations, not the modified rule that required one control observation to be at least two within-run standard deviations above its QC mean and the other be at least 2 within-run standard deviations below its mean. Unlike the 1_{3s} and 2_{2s} rules, the rejection region for the

Table 1. Probabilities of Rejection Obtained by Numerical Integration

		Probability of rejection[a]			
Rule	ϕ_s[b]	In-control	SE = 2.0	RE_b = 1.5	RE_w = 1.5
1_{3s}	0	0.0054	0.2921	0.0814	0.0889
1_{3s}	1	0.0052	0.2548	0.0726	0.0878
2_{2s}	0	0.0010	0.2500	0.0634	0.0166
2_{2s}	1	0.0081	0.3333	0.0966	0.0310
R_{4s}	0	0.0047	0.0047	0.0047	0.0593
R_{4s}	1	0.0047	0.0047	0.0047	0.1306
$\bar{X}$	0	0.0050	0.5085	0.1335	0.0613
$\bar{X}$	1	0.0050	0.3094	0.0856	0.0382
$1_{3s}/2_{2s}/R_{4s}$	0	0.0097	0.4089	0.1117	0.1234
$1_{3s}/2_{2s}/R_{4s}$	1	0.0159	0.3971	0.1204	0.1900
$\bar{X}/R_{4s}$	0	0.0100	0.5194	0.1406	0.1189
$\bar{X}/R_{4s}$	1	0.0100	0.3203	0.0923	0.1651

[a] SE denotes a shift in the QC mean; RE_b and RE_w denote increases in total analytical imprecision due to an increase in the between-run component and within-run component, respectively.

[b] Ratio of between-run to within-run imprecision during stable operation.

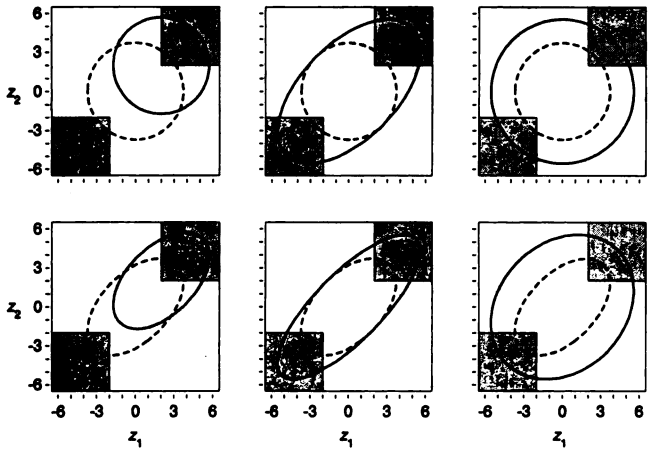

Fig. 3. Same as in Fig. 2, but representing the error detection characteristics of the 2_{2s} rule

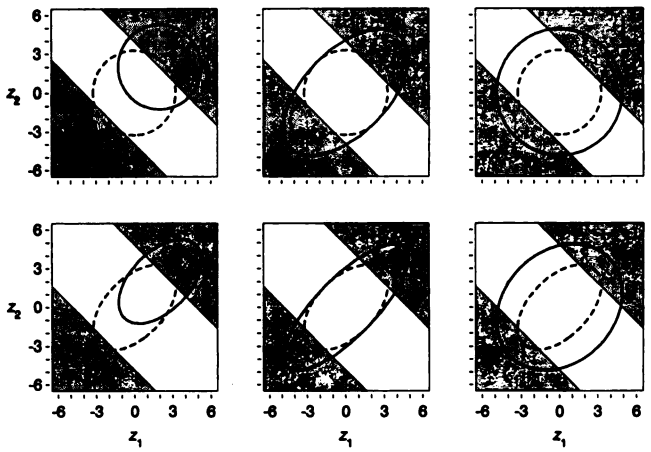

Fig. 5. Same as in Fig. 2, but representing the error detection characteristics of the $\bar{X}$ rule

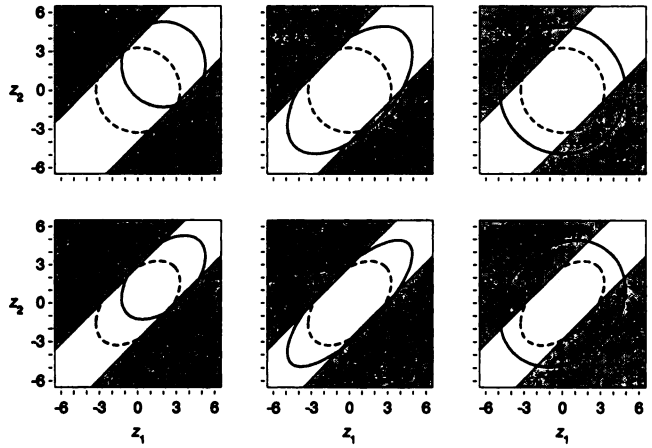

Fig. 4. Same as in Fig. 2, but representing the error detection characteristics of the R_{4s} rule

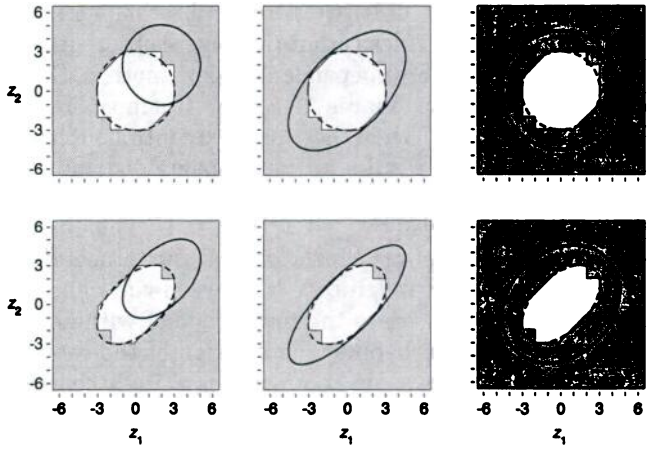

Fig. 6. Same as in Fig. 2, but representing the error detection characteristics of the $1_{3s}/2_{2s}/R_{4s}$ rule

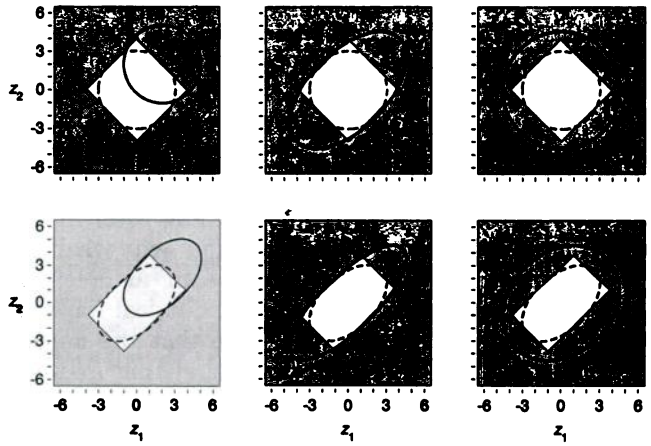

Fig. 7. Same as in Fig. 2, but representing the error detection characteristics of the $\bar{X}/R_{4s}$ rule

R_{4s} rule changes as the ratio of between-run to within-run imprecision changes. This is because z_1 and z_2 are defined in terms of total analytical standard deviation, but the R_{4s} rule is defined in terms of within-run standard deviation. The *Appendix* describes how the rejection regions in Figure 4 are determined.

Figure 5 depicts the $\bar{X}$ rule. This rule rejects a run if the absolute value of the average of z_1 and z_2 exceeds the specified control limit. The $\bar{X}$ rule shown in Figure 5 has a control limit of 2.807, which gives a false-rejection rate of 0.005. The variance of the average of z_1 and z_2 equals $(1 + \rho)/2$. The rejection region for the $\bar{X}$ rule also changes as the ratio of between-run to within-run imprecision changes (as in the R_{4s} rule), because ρ depends on the relative amounts of between-run and within-run imprecision. The *Appendix* describes how the rejection regions in Figure 5 are determined.

Figure 6 depicts the rejection region resulting from the combination of the 1_{3s}, 2_{2s}, and R_{4s} rules, and Figure 7 portrays the combination of the $\bar{X}$ and R_{4s} rules. Control limits of 2.785 were used for the $\bar{X}$ rule in Figure 7. When combined with the R_{4s} rule, this produces a false-rejection rate of 0.01. The shaded rejection region in Figure 6 is the result of overlaying the shaded regions for each individual control rule. Consequently,

the rejection probability for this combination rule is greater than for any of the three individual rules, but (because the individual rejection regions overlap) less than the sum of the rejection probabilities of the three individual rules. Also in Figure 6, the amount of shaded region contained within the dashed contour is greater in the three lower panels, where the between-run and within-run components of imprecision are equal, than in the three upper panels, where the between-run component of imprecision is zero. This shows that the false-rejection rate for the $1_{3s}/2_{2s}/R_{4s}$ rule increases as ϕ_s increases. Table 1 demonstrates the effect of ϕ_s on the various control rules. The rejection regions for the R_{4s} and $\bar{X}$ rules change as ϕ_s changes so that the false-rejection rate remains unchanged. However, the rejection rates for the 1_{3s} and 2_{2s} rules change as ϕ_s changes. Consequently, the rejection rate of the $1_{3s}/2_{2s}/R_{4s}$ rule depends on ϕ_s but that of the $\bar{X}/R_{4s}$ rule does not.

Discussion

One of the effects of between-run imprecision on the performance of QC rules is that the control observations within a run are not independent of one another. Each control observation shares a common component of between-run error. When there are two control observations per run, this results in the bivariate distribution for the control observations being more concentrated along the line defined by $z_2 = z_1$ (Figure 1). This line is referred to as the principal axis of the probability distribution. As the magnitude of between-run imprecision increases relative to the magnitude of within-run imprecision, the probability distribution of the control observations within a run becomes increasingly concentrated along the principal axis.

Visualizing the consequences that the three different error conditions studied here have on the probability distribution of the control observations helps in the interpretation of the performance of the various QC rules. The distribution moves along the principal axis when an out-of-control condition causes equal shifts (relative to total analytical imprecision) in each QC mean. The distribution is "stretched" along the principal axis when an out-of-control condition in between-run imprecision causes equal increases in the total analytical imprecision for each control material. The distribution is spread out in all directions, but to a greater degree (when ϕ_s is greater than zero) in the direction perpendicular to the principal axis when an out-of-control condition in within-run imprecision causes equal increases in the total analytical imprecision for each control material. Note that the middle panels of Figures 2–7 illustrate that, within a single run, increases in between-run imprecision behave more like a shift in the mean than an increase in analytical imprecision (*3*).

Even though many more-sophisticated QC rules have been proposed, the 1_{3s} rule is still widely used in clinical laboratories. Figure 2 verifies that the 1_{3s} rule is a reasonable all-around rule that protects against all three types of error conditions, because it guards against changes in the probability distribution of the control observations in all directions. In contrast, the 2_{2s} rule (Figure 3) can detect only changes in the distribution along its principal axis. This illustrates why it is much better at detecting shifts in the mean or increases in between-run imprecision than at detecting increases in within-run imprecision. The R_{4s} rule (Figure 4) detects only changes perpendicular to the principal axis of the probability distribution, which explains why it has power for detecting increases in within-run imprecision but not shifts in the mean or increases in between-run imprecision.

The $\bar{X}$ rule (Figure 5), like the 2_{2s} rule, detects changes along the principal axis of the probability distribution. However, the $\bar{X}$ rule maximizes the amount of the probability distribution falling within the rejection region whenever an error condition causes a change in the distribution along its principal axis, because this rule's rejection limits are perpendicular to the principal axis. Thus, the rule demonstrates good power at detecting the types of shifts in the QC means and increases in between-run imprecision evaluated here (*3, 11*). In fact, it can be proved that of all rules based on control observations within a single run with the same false-rejection rate, the $\bar{X}$ rule is the most powerful for detecting these out-of-control conditions (*12*).

Figure 6 clearly illustrates that the impact of combining the 1_{3s}, 2_{2s}, and R_{4s} rules is to create a rejection region that more closely approximates the contour of the in-control probability distribution of the control observations. This produces a rule that has increased ability to detect all types of error conditions (but with an accompanying increase in false-rejection rate). Another result that can be theoretically derived is that the QC rule that uses the dashed contours in the figures to define its rejection region is the most powerful rule for detecting the increases in within-run imprecision studied here (*12*). The quadratic equation that defines these contours of constant probability density represents a chi-square statistic (χ^2). Therefore, the $1_{3s}/2_{2s}/R_{4s}$ multirule approximates the χ^2 rule, which is the best rule for detecting these increases in within-run imprecision. The *Appendix* shows that this χ^2 rule is a combination of the quantities used in the $\bar{X}$ rule and the range rule.

The $\bar{X}/R_{4s}$ rule (Figure 7) defines a rejection region that is similar in appearance to the rejection region for the 1_{3s} rule, with two important exceptions. First, unlike the 1_{3s} rule, the rejection region for the $\bar{X}/R_{4s}$ rule changes with ϕ_s so that the probability of false rejection does not depend on the value of ϕ_s. It is not readily apparent from Figure 2 and Table 1, but p_{fr} and p_{ed} for the 1_{3s} rule both decrease as ϕ_s increases. The false-rejection rate for the 1_{3s} rule when $\phi_s = 2$ is 0.0047 and when $\phi_s = 4$ it is 0.0039. Second, unlike the 1_{3s} rule, the orientation of the rejection regions for the $\bar{X}/R_{4s}$ rule is optimal with respect to the changes in the probability distribution of the control observations that occur during the three types of out-of-control conditions investigated here.

Clearly, no one rule is best under all out-of-control error conditions. If detecting all types of error conditions is important, a compromise is required. Table 2 displays rejection probabilities for the $\bar{X}$ rule, the χ^2 rule, the $\bar{X}/R_{4s}$ rule, and the $1_{3s}/2_{2s}/R_{4s}$ rule for the three error conditions, with the control limits of the rules defined so their false rejection rates match the false rejection rate of the $1_{3s}/2_{2s}/R_{4s}$ rule. Table 2 suggests that the $\bar{X}/R_{4s}$ rule is a good compromise. Interestingly, this is the same QC strategy proposed by Levey and Jennings (*13*) in one of the earliest papers on quality control to appear in the laboratory medicine literature. More recently, Linnet (*8*) provided strong arguments supporting the superiority of the mean rule and a rule based on the within-run standard deviation, s_w^2. The s_w^2 rule has been described previously where it was denoted $s_{w,1}^2$ (*3*) or s_{wj}^2 (*11*). With two control observations per run, the s_w^2 rule and the range rule are equivalent (*3*).

The χ^2 rule also displays respectable power in all cases, with slightly superior performance in detecting increases in within-run imprecision but slightly inferior performance in detecting shifts in the QC means or increases in between-run imprecision compared with the $\bar{X}/R_{4s}$ rule. However, the $\bar{X}/R_{ks}$ combination rule (where k can be any positive value, including 4) allows two control limits to be specified. This provides flexibility in defining the relative error detection ability for shifts in the QC means and increases in between-run imprecision vs increases in within-run imprecision for cases in which the importance of detecting these different types of error conditions vary. The single control limit of the χ^2 rule does not allow control over the relative error detection ability for the different types of out-of-control conditions.

The results displayed in Tables 1 and 2 and Figures 1–7 reflect only a single error magnitude for each out-of-control condition. The same relative relation between the control rules holds at other error magnitudes, although the degree of difference between the rules changes as the magnitude of error changes.

If the two control observations within an analytical run are on a single control material, any out-of-control condition will necessarily produce the same shift in the means of z_1 and z_2 or increase in the total analytical imprecision of z_1 and z_2. However, if the two control observations being evaluated by the control procedure are on different control materials, then out-of-control conditions are not restricted to equal shifts in the means or increases in total imprecision. The consequence of an out-of-control condition that produces unequal shifts in the QC means of the two control materials is to move the bivariate distribution of z_1 and z_2 in a different direction than the principal axis. For example, an out-of-control condition that causes a relatively large shift in the QC mean for a low-concentration control material (z_1), but a negligible shift in the mean for a high-concentration control material (z_2) would move the bivariate distribution of z_1 and z_2 approximately along the direction of the z_1 axis. Likewise, an out-of-control condition that causes unequal increases in total imprecision for the two control materials will expand the bivariate distribution of z_1 and z_2 and will also rotate the distribution so that its principal axis is no longer defined by the line $z_2 = z_1$. The laboratory medicine QC literature has investigated only out-of-control conditions that produce equal shifts in the QC means or equal increases in imprecision.

When out-of-control conditions can produce changes in the distribution of control observations in any direction, the situation more appropriately becomes a problem in multivariate QC (*14, 15*). If it is important to detect any shifts in the QC means or any increases in total imprecision, then the χ^2 rule that defines the in-control contour for the control observations has optimal characteristics. This was briefly mentioned by Heilbron et al. (*7*). However, if out-of-control conditions are expected to produce similar shifts in each QC mean or similar increases in total imprecision for each control material, then the $\bar{X}/R_{ks}$ rule is again a good compromise.

When the number of control observations in a run (N) is greater than two, displaying the error detection characteristics of the QC rules in the fashion employed here becomes impractical. With $N = 3$, the acceptance region for the 1_{3s} rule and the rejection regions for the 2_{2s} rule are cubes in three-dimensional space. The rejection boundary for the R_{4s} rule is a cylinder aligned along the principal axis with a hexagonal cross-section. The $\bar{X}$ rule's rejection limits are two parallel planes perpendicular to the principal axis. However, the conclusions are the same. The $\bar{X}$ rule is still best for detecting shifts of equal magnitude (relative to total imprecision) in the QC means or increases in between-run imprecision that produce equal increases in total imprecision for each control material. The superiority of the $\bar{X}$ rule over the other rules increases as N increases. The χ^2 rule representing the probability contour for the joint distribution of the N control observations is still best for detecting increases in within-run imprecision that produce equal increases in total imprecision for each control material. With $N > 2$, the s_w^2 rule is superior to the range rule. The rejection limits of the s_w^2 rule

Table 2. Probabilities of Rejection after Matching False-Rejection Rates

		Probability of rejection[a]			
Rule	ϕ_s[b]	In-control	SE = 2.0	RE_b = 1.5	RE_w = 1.5
$\bar{X}$	0	0.0097	0.5953[c]	0.1667[c]	0.0846
$\bar{X}$	1	0.0159	0.4592[c]	0.1397[c]	0.0749
χ^2	0	0.0097	0.4827	0.1316	0.1273[c]
χ^2	1	0.0159	0.3555	0.1067	0.2082[c]
$\bar{X}/R_{4s}$	0	0.0097	0.5114	0.1378	0.1171
$\bar{X}/R_{4s}$	1	0.0159	0.4135	0.1247	0.1838
$1_{3s}/2_{2s}/R_{4s}$	0	0.0097	0.4089	0.1117	0.1234
$1_{3s}/2_{2s}/R_{4s}$	1	0.0159	0.3971	0.1204	0.1900

[a] SE denotes a shift in the QC mean; RE_b and RE_w denote increases in total analytical imprecision due to an increase in between-run and within-run components, respectively.

[b] Ratio of between-run to within-run imprecision during stable operation.

[c] The maximum obtainable probability of rejection for the error condition by any rule with the same false-rejection rate.

when $N = 3$, for example, define a cylinder aligned along the principal axis with a more optimal circular (rather than hexagonal) cross-section. Therefore, the $\bar{X}/s_w^2$ rule is superior to the $\bar{X}/R_{ks}$ rule. However, for values up to $N = 6$ it has been demonstrated that the difference between the s_w^2 rule and the R_{ks} rule is small (*3*).

In summary, important conclusions concerning the error detection characteristics of QC rules that use control observations within a single analytical run have been reached without performing simulations. With the assumption of two control observations per run, all false-rejection and error detection probabilities have been determined exactly by numerical integration. Graphical display of the probability distributions of the two control observations in a run under various in-control and out-of-control conditions superimposed on the rejection region of a QC rule has been used to lend insight into the comparative performance characteristics of alternative control strategies. The numerical results and visual insight support the conclusion that no single QC rule is best under all error conditions, but the $\bar{X}/s_w^2$ (or $\bar{X}/R_{ks}$) rule offers a good compromise.

Appendix

Rejection Limits for the R_{4s} Rule Shown in Figure 4

The rejection limits for the R_{4s} rule are defined as

$$[(C_1 - \mu_1)/\sigma_{w_1}] - [(C_2 - \mu_2)/\sigma_{w_2}] = \pm 4 \quad (1)$$

or

$$(z_1\sigma_{t_1}/\sigma_{w_1}) - (z_2\sigma_{t_2}/\sigma_{w_2}) = \pm 4 \quad (2)$$

where $z_j = (C_j - \mu_j)/\sigma_{t_j}$. Given that $\phi_{s_j} = \sigma_{b_j}/\sigma_{w_j}$ and $\sigma_{t_j}^2 = \sigma_{b_j}^2 + \sigma_{w_j}^2$, it follows that $\sigma_{t_j}/\sigma_{w_j} = \sqrt{1 + \phi_{s_j}^2}$. If $\phi_{s_1} = \phi_{s_2} = \phi_s$, then $\rho = \phi_s^2/(1 + \phi_s^2)$ and $\sigma_{t_j}/\sigma_{w_j} = 1/\sqrt{1 - \rho}$. Upon substitution, the rejection limits become

$$(z_1 - z_2)/\sqrt{1 - \rho} = \pm 4 \quad (3)$$

or

$$z_2 = z_1 \pm 4/\sqrt{1 - \rho} \quad (4)$$

If $\phi_{s_1} = \phi_{s_2} = 0$, then $\rho = 0$ and the rejection limits are defined by the two lines $z_2 = z_1 \pm 4$. If $\phi_{s_1} = \phi_{s_2} = 1$, then $\rho = 0.5$ and the rejection limits are defined by the lines $z_2 = z_1 \pm 4/\sqrt{2}$.

Rejection Limits for the $\bar{X}$ Rule Shown in Figure 5

The variance of $(z_1 + z_2)/2$ equals $\{\mathrm{Var}(z_1) + \mathrm{Var}(z_2) + 2\mathrm{Cov}(z_1, z_2)\}/4$ where Var and Cov represent variance and covariance, respectively. The variances of z_1 and z_2 are 1 and their covariance is ρ. Therefore, the variance of $(z_1 + z_2)/2$ is $\{1 + 1 + 2\rho\}/4 = (1 + \rho)/2$. The rejection regions depicted in Figure 5 are of the form

$$\frac{(z_1 + z_2)/2}{\sqrt{(1 + \rho)/2}} = \pm 2.807 \quad (5)$$

or

$$z_2 = -z_1 \pm 2.807\sqrt{2(1 + \rho)} \quad (6)$$

If $\phi_{s_1} = \phi_{s_2} = 0$, then $\rho = 0$ and the rejection limits are $z_2 = -z_1 \pm 2.807\sqrt{2}$. If $\phi_{s_1} = \phi_{s_2} = 1$, then $\rho = 0.5$ and the rejection limits are $z_2 = -z_1 \pm 2.807\sqrt{3}$.

Rejection Limits Derived from the Bivariate Probability Contours

Assume z_1 and z_2 follow a bivariate Gaussian distribution with means = 0, variances = 1, and correlation = ρ. Then the equation that defines the contour of constant elevation on the density surface at a height = h is (*16*, p 89)

$$\frac{z_1^2 - 2\rho z_1 z_2 + z_2^2}{1 - \rho^2} = h \quad (7)$$

The numerator and denominator can be reexpressed as

$$\frac{(z_1 - z_2)^2(1 + \rho)/2 + (z_1 + z_2)^2(1 - \rho)/2}{(1 + \rho)(1 - \rho)} = h \quad (8)$$

which simplifies to

$$\frac{(z_1 - z_2)^2}{2(1 - \rho)} + \frac{(z_1 + z_2)^2}{2(1 + \rho)} = h \quad (9)$$

The first term in this sum is proportional to the square of the statistic used in the range rule and the second term is proportional to the square of the statistic used in the mean rule (see above). Each of these terms is chi-square distributed with one degree of freedom and they are statistically independent. Therefore, their sum is a chi-square statistic with two degrees of freedom, and the control limit h can be specified as an appropriate percentile of the distribution.

References

1. Parvin CA. Comparing the power of quality-control rules to detect persistent systematic error. Clin Chem 1992;38:358–63.
2. Parvin CA. Estimating the performance characteristics of quality-control procedures when error persists until detection. Clin Chem 1991;37:1720–4.
3. Parvin CA. Comparing the power of quality-control rules to detect persistent increases in random error. Clin Chem 1992;38:364–9.
4. Westgard JO, Groth T. Design and evaluation of statistical control procedures: applications of a computer "quality control simulator" program. Clin Chem 1981;27:1536–45.
5. Hatjimihail AT. A tool for the design and evaluation of alternative quality-control procedures. Clin Chem 1992;38:204–10.
6. Westgard JO. Simulation and modeling for optimizing quality control and improving analytical quality assurance. Clin Chem 1992;38:175–8.
7. Heilbron DC, Eastman JW, Kelly D. Warning methods for quality control in determinations of serum calcium. Clin Chem 1974;20:1416–21.

8. Linnet K. Mean and variance rules are more powerful or selective than quality control rules based on individual values. Eur J Clin Chem Clin Biochem 1991;29:417–24.
9. Westgard JO, Groth T, Aronsson T, Falk H, de Verdier C-H. Performance characteristics of rules for internal quality control: probabilities for false rejection and error detection. Clin Chem 1977;23:1857–67.
10. Wolfram S. Mathematica: a system for doing mathematics by computer, 2nd ed. Redwood City, CA: Addison-Wesley, 1991.
11. Linnet K. The between-run component of variation in internal quality control. Clin Chem 1989;35:1416–22.
12. Kendall M, Stuart A. The advanced theory of statistics, Vol. 2, 4th ed. New York: Macmillan, 1979.
13. Levey S, Jennings ER. The use of control charts in the clinical laboratories. Am J Clin Pathol 1950;20:1059–66.
14. Alt FB. Multivariate quality control. In: Kotz S, Johnson NL, Read CR, eds. The encyclopedia of statistical sciences. New York: John Wiley, 1984:110–22.
15. Jackson JE. Multivariate quality control. Commun Stat Theor Methods 1985;14:2657–88.
16. Morrison DF. Multivariate statistical methods, 2nd ed. New York: McGraw-Hill, 1976.