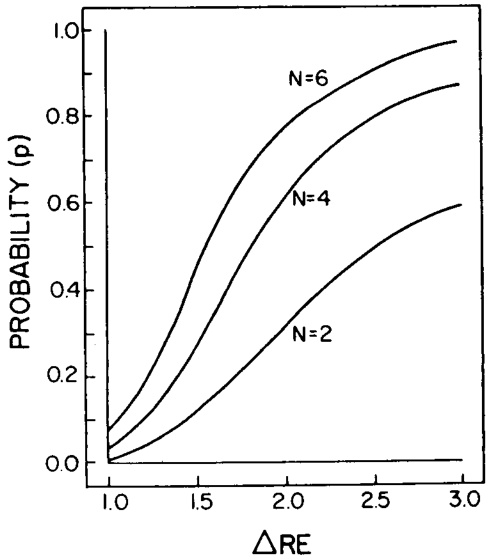

Fig. 4. Power of the $1_{3s}/2_{2s}/R_{4s}/4_{1s}/10_{\bar{x}}$ multi-rule Shewhart procedure for detecting increases in random error

The probability for rejection (*p*) is plotted on the ordinate vs size of the random error (ΔRE) on the abscissa. A value of 2.0 on the *x*-axis refers to a doubling of the standard deviation of the analytical method

of the R_{4s} rule on day 11 in Figure 3, this is usually related to different causes, such as the instability of reagents or measurement conditions, or variability in timing, pipetting, individual technique, or other similar factors. Definition of the possible sources of errors depends on the particular analytical method and the instrumentation used. The analyst will be assisted by the manufacturer's troubleshooting guidelines, documentation of reagent or instrument changes, documentation of previous problems, and experience.

When a control problem has been resolved, there remains a question of what should be done with the control data from that run—whether it should be included in further assessment of control status and in further data calculations. In assessing control after problem-solving procedures have been used, the objective should be to assess control of the newly corrected process. This is best done by increasing the number of control observations in that next run, rather than utilizing any observations from a previous run. In performing calculations on control data to update the control limits, the purpose is to characterize only the *stable* performance of the analytical process. Data obtained during unstable operation should not be included.

Note: Reviewer R.B. emphasizes the importance of defining how control results from out-of-control runs should be handled. The authors' perspective is that these results do not represent the stable performance of the analytical method; therefore, if included in the summary data calculations, they would cause the standard deviation to be too large and the resulting control limits too wide. On the other hand, there is concern that elimination of these points will narrow the control limits, so that they no longer correctly characterize the tails of the error distribution. This latter problem should be minimized here because observations between $2s$ and $3s$ will be included in the final summary calculations.

Performance Characteristics

The performance characteristics of this multi-rule Shewhart procedure are summarized by the "power functions" given in Figures 4 and 5. These plots show the probability for rejecting an analytical run as a function of the size of the error occurring in the run (*5*). The probability for rejection (*p*) is plotted in the *y*-direction vs the size of the analytical error (ΔRE, ΔSE) in the *x*-direction. The point of intersection on the *y*-axis gives the probability for false rejection (the probability for rejecting the run when there is no analytical error except for the inherent imprecision of the analytical method). Points on the curves give the probability for error detection (the probability for rejecting the run when there is an error of the size indicated on the *x*-axis). In Figure 4, the size of the random error (ΔRE) is given as a factor such that a value of 2.0 means that the standard deviation of the method has doubled. In Figure 5, the size of the systematic error (ΔSE) is given in multiples of the standard deviation, such that a value of $2.0s$ means a systematic error equivalent to two times the standard deviation of the analytical method.

Note: Reviewer A.H. stressed the importance of understanding the concept of *inherent* random error. There is always some random error associated with a measurement process, even a stable and well-controlled process. When we discuss error detection, we actually refer to detecting error which is in addition to that inherent random error. This concept of error is important, because it helps explain why it is difficult for control procedures to detect small analytical errors. In effect, we are dealing with a signal-to-noise problem, with the inherent random error being the noise, and the additional analytical error being the signal we would like to detect.

Note: Reviewer R.T. commented that he had performed some simulations on these procedures also and that they revealed these procedures to be valid.

The different lines in Figures 4 and 5 represent the number of observations per analytical run (N). As N increases, the capability for error detection increases. However, when N = 6, the probability for false rejection exceeds 5%. This is primarily ascribable to the R_{4s} rule, and elimination of that rule will decrease the false rejections to an acceptable proportion. Based on the performance characteristics shown in Figures 4 and 5, the control procedure recommended here should be satisfactory for N from 2 to 4. If N is greater than 4, then the R_{4s} rule should be eliminated, or use of other control procedures should be considered.

Power functions should provide the basis for comparing the performance of different control procedures. However, such information is often lacking in the references and descriptions of other procedures, often because the data testing and interpretation are not sufficiently well defined. Very careful and detailed guidelines are necessary in order adequately to de-

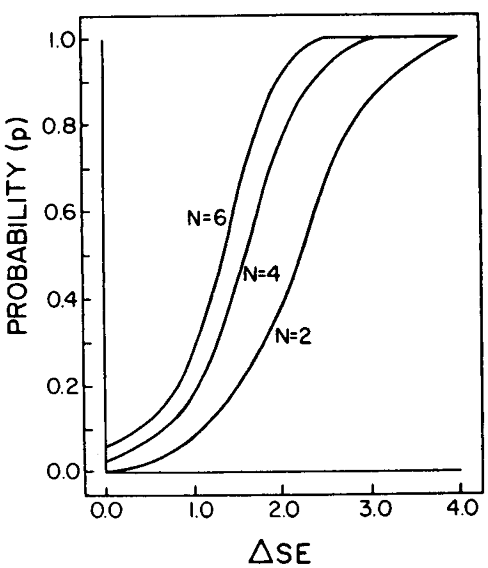

Fig. 5. Power of the $1_{3s}/2_{2s}/R_{4s}/4_{1s}/10_{\bar{x}}$ multi-rule Shewhart procedure for detecting systematic error

The probability for rejection (*p*) is plotted on the ordinate vs the size of the systematic error (ΔSE) on the abscissa. A value of 2.0 on the *x*-axis refers to a systematic shift equivalent to two times the standard deviation of the analytical method