

where $\hat{x}(K)$ =predicted state (e.g., concentration) at time K , $\bar{x}(K-1)$ =the estimation of x at $K-1$, F =the transition function, and $w(K)$ =the variance of noise.

The next step is to do a new measurement $z(K)$.

$$\bar{x}(K) = \hat{x}(K) + G(K)[\hat{x}(K) - H(K)z(K)],$$

where $H(K)$ =calibration factor.

The estimate $x(K)$ is the best estimate that is possible. The weighting factor

$$G(K) = P(K)[P(K) + R(K)]^{-1}$$

depends on the fluctuations in the system, expressed as covariance $P(K)$ and the measuring error $R(K)$.

A Kalman filter enables the on-line estimation of calibration parameters, intercept, sensitivity and drift (of both intercept and sensitivity). The filter requires a model of the system, including system noise and measurement noise. When a good model is available, the filter can predict future values or estimate best values of the changing parameters. These figures may be used to determine when a recalibration is required.

Making a number of assumptions, the usual Kalman filter algorithms can be used resulting in [4,5,6]

$$\begin{aligned}\hat{x}(K) &= F(K)\bar{x}(K-1) \\ \hat{P}(K) &= F(K)\hat{P}(K-1)F'(K) + Q(K-1) \\ \bar{x}(K) &= \hat{x}(K) + G(K)\{z(K) - h'(K)\hat{x}(K)\} \\ \bar{P}(K) &= \hat{P}(K) - G(K)h'(K)\hat{P}(K) \\ G(K) &= \hat{P}(K)h(K)\{h'(K)\hat{P}(K)h(K) + R(K)\}^{-1}.\end{aligned}$$

The quality performance of the system is controlled on-line by comparison with a predefined criterion. The decision to re-calibrate is followed by optimization of the concentration standards available. The example given treats drift as a deterministic phenomenon. Poulisse [7,8] indicated another approach, considering drift as a stochastic phenomenon. A novel approach aiming at the elimination of the human factor in calibration is the use of expert systems. An expert system allows the objective use of a large knowledge base. This knowledge base consists of facts and rules, the results of long experience. A so-called inference machine searches for the right solution given a set of starting parameters or states the starting parameters when the goal has been stated. Since the computer will not be annoyed when repeatedly asked for the conditions for a good calibration and can deliver

calibration correction factors for each situation, it can be expected that this approach will give better calibration results. However, no conclusive results have been reported so far.

References

- [1] Kalivas, J. H., and Kowalski, B. R., *Anal. Chem.* **54**, 560 (1982).
- [2] Smit, H. C., Mars, C., and Kraak, J. C., *Anal. Chim. Acta* **181**, 37 (1986).
- [3] Müskens, P. J. W. M., *Anal. Chim. Acta* **103**, 445 (1978).
- [4] Thijssen, P. C., Wolfrum, S. M., Smit, H. C., and Kateman, G., *Anal. Chim. Acta* **156**, 87 (1984).
- [5] Thijssen, P. C., Smit, H. C., and Kateman, G., *Anal. Chim. Acta* **162**, 253 (1984).
- [6] Poulisse, H. N. J., and Engelen, P., *Anal. Letters* **13** (A14), 1211 (1980).
- [7] Poulisse, H. N. J., "State- and parameter estimation as chemometrical concepts," thesis, Nijmegen 1983, chapter 6.
- [8] Poulisse, H. N. J., "State- and parameter estimation as chemometrical concepts," thesis, Nijmegen 1983, chapter 8.

Design of Cost-Effective QC Procedures for Clinical Chemistry Assays

James O. Westgard

Department of Pathology and
Laboratory Medicine
Medical School, Medical Technology Program
School of Allied Health Professions
University of Wisconsin
Madison, WI

Introduction

In clinical laboratories, the major production processes are analytical processes. Quality management strategies must include approaches for optimizing the cost-effectiveness of analytical processes. "Cost-effective QC," in this context, is concerned with selecting or designing a quality control procedure that maximizes both the quality and productivity of an analytical process [1]. There are many factors that need to be considered, including the medically required quality for the analyte being measured, the characteristics of the measurement procedure (type of process or system, precision,

accuracy, drift stability, and frequency of occurrence of errors), the type and structure of the errors occurring (random or systematic, intermittent or persistent), and the error detection and false rejection characteristics of the control procedure itself (probabilities for rejection, average run lengths).

To assess the possible interactions of these many factors, it is necessary to have a "design tool" such as a quality-productivity model, which is based on the industrial concept of quality-costs. With the aid of such models, the effects of these many factors can be predicted, facilitating the planning of analytical processes that will provide cost-effective operation.

Quality-Productivity Models

The industrial concept of "quality costs" was described by Feigenbaum in 1956 [2]. Quality costs include components of prevention costs, appraisal costs, and failure costs, including both internal and external failure costs. When costs are restricted to process costs in terms of samples analyzed or repeated, prevention costs are losses of samples for calibration, and appraisal costs are losses for analyzing control samples, both of which can be estimated from knowledge of the physical variables of the analytical system and process. Failure costs are losses for repeat runs and repeat requests, which can be predicted from performance characteristics of the measurement and control procedures. These losses, described in terms of the process variables and characteristics, provide an estimate of process waste. Test yield, or productivity, is estimated by the difference between the maximum output and the process waste ($1 - [\text{average quality costs or losses}]$).

Table 1 provides an example quality-productivity model for a batch analytical process subject to intermittent errors.

Table 1. A quality-productivity model for describing the quality and productivity of a batch analytical process subject to intermittent analytical errors

Measure of quality is "defect rate"

$$\text{Defect Rate} = f(1 - P_{ed})$$

Measure of productivity is "test yield"

$$\text{Test Yield} = 1 - \frac{C+N}{T} - \frac{Sp}{T} [R_{re}fP_{ed} + R_{fr}(1-f)P_{fr} + R_{ra}f(1-P_{ed}) + R_{ra}f(1-P_{ed})(1-f)(1-P_{fr})]$$

where the terms are characteristics of the measurement and control procedures, and laboratory policies for repeat runs

- f is the frequency of occurrence of errors
- P_{ed} is the probability for error detection
- P_{fr} is the probability for false rejection
- C is the number of calibrators
- N is the number of control measurements per run
- Sp is the number of patient specimens
- T is the total samples in a run ($C + N + Sp$)
- R terms are rerun factors for true reject, false reject, false accept, and true accept runs

Example Application

Let us assume a batch process having 30 samples per batch that is calibrated once a week, requiring six calibrators (C) be analyzed in duplicate. Assume a workload requiring one run per day for 6 days in a week, thus (C) averages one per run for this example. We then postulate a design for a control procedure, enter the number of control measurements per run (N), and the probabilities for false rejection and error detection (for medically important errors), as determined from power function graphs obtained by a computer simulation program [3]. Complete details of this example are found elsewhere [1].

Figure 1 shows the performance of a multi-rule procedure [4] for $N=2$ and $N=4$. For an unstable measurement procedure having a frequency of errors of 0.10 (or 10%), doing more quality control actually provides higher quality and higher productivity. Doubling the number of controls per run actually reduces the cost. For a stable measurement procedure, with a frequency of errors of 0.01 (or 1%), higher productivity is achieved with $N=2$.

Figure 2 compares the performance of three different control procedures having $N=2$. For a stable measurement procedure, the use of 3s control limits (1₃, control rule) or a multi-rule control pro-

cedure [3] provides high quality (low defect rate) and high productivity (high test yield). For an unstable measurement procedure, say a frequency of errors of 0.20 (or 20%), the use of 2s control limits (1_{2s} control rule) provides better quality and higher productivity.

Figure 3 shows a comparison of three different control procedures, with different rules and with different N 's, all giving nearly the same quality or defect rates. Multi-rule with $N=4$ or 1_{2s} with $N=2$ provide more cost-effective operation than 1_{3s} with $N=8$.

Implications for Selecting and Designing QC Procedures

Given the practical limitations of two to four control measurements per run due to short runs for fast turnaround of results, it is impossible to have an ideal control procedure with *both* high error detection and low false rejection. However, a laboratory can have *either* high error detection or low false rejection if it knows which is more important for the particular application.

Because the performance of an analytical process depends on its stability, i.e., its frequency of errors (f), the control procedure can be tailored to fit that frequency of errors. For measurement procedures having a low frequency of errors, the QC procedure should be selected or designed primarily for a low probability of false rejection, with secondary concern for error detection ("low- f design"). When there is a high frequency of errors, the QC procedure should be selected or designed for high error detection, with little regard for its probability for false rejection ("high- f design").

One useful strategy is to change control procedures depending on the recent history of the measurement procedure, switching from "low- f " to "high- f " designs, as necessary. Whenever the analyst is suspicious that performance is deteriorating, the "high- f " design can be used. When the process is running smoothly, the "low- f " design can be used.

Another strategy is to systematically switch from one design to another as the analytical process goes through its normal operational cycle. For example, at the beginning of the day, test the process carefully using a "start-up" design having high error detection; following successful completion of that testing, switch to a "monitoring" design having very low false rejection; and, over the course

of the day, use a "prospective" design to analyze the accumulated control measurements and trigger preventive maintenance procedures prior to tomorrow's run. Such "multi-stage" designs provide a systematic tailoring of the control procedure to fit the changing operation conditions of the measurement procedure.

Doing "cost-effective QC" in clinical laboratories means using individualized designs of QC procedures that fit the measurement procedures to be monitored. A comprehensive design tool would be useful to help users study and plan their analytical processes. It should provide a wide variety of control rules, incorporate computer simulation to provide the performance characteristics of the selected QC procedure, utilize more realistic error distributions to represent stable and unstable performance, offer choices of intermittent and/or persistent error conditions, and permit input of information on the factors and variables needed to customize the quality-productivity model to fit individual analytical systems.

Routine implementation of such individualized QC designs also would be aided by more advanced QC programs in laboratory information systems. The programs should permit analysts to choose the control rules and the number of control measurements, to select how the rules are applied across materials and/or runs, to specify whether the control signal is a rejection or warning for preventive maintenance, and to define two or more QC procedures that could be employed on a single measurement procedure.

Accuracy in Trace Analysis

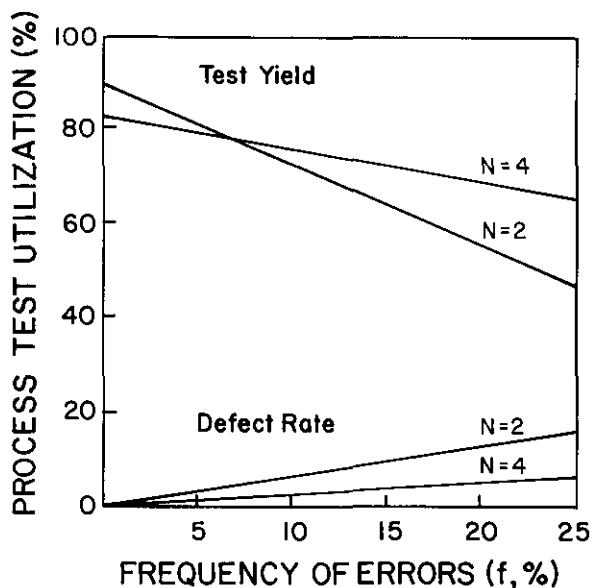


Figure 1. Comparison of quality (defect rate) and productivity (test yield) for a batch analytical process using a multi-rule control procedure with $N=2$ and $N=4$. From reference [1].

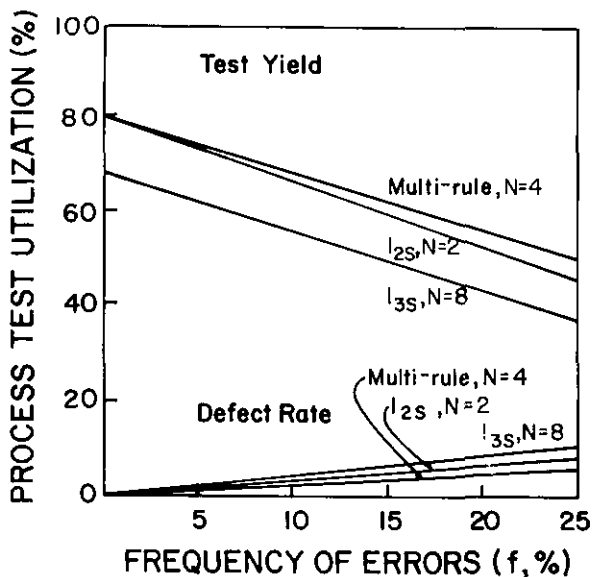


Figure 3. Comparison of quality (defect rate) and productivity (test yield) for a batch analytical process with different control rules and different N 's. From reference [1].

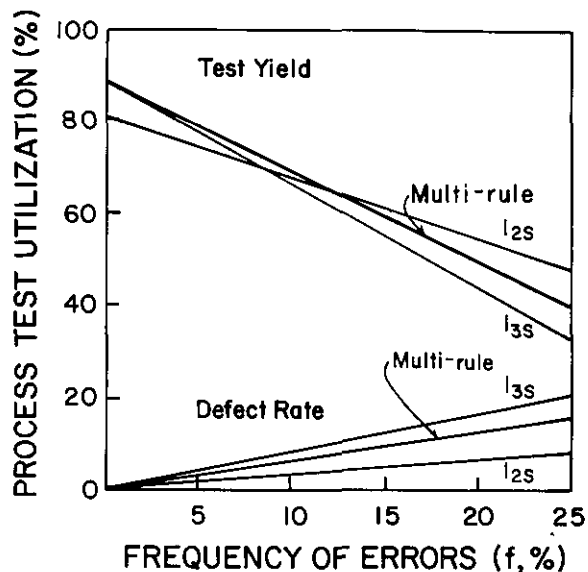


Figure 2. Comparison of quality (defect rate) and productivity (test yield) for a batch analytical process with different control rules all using $N=2$ (2 control measurements per run). From reference [1].

References

- [1] Westgard, J. O., and Barry, P. L., *Cost-Effective Quality Control*, AACC Press, Washington, DC (1986).
- [2] Feigenbaum, A. V., *Harvard Bus. Rev.* 34, 93 (1956).
- [3] Westgard, J. O., and Groth, T., *Clin. Chem.* 25, 863 (1979).
- [4] Westgard, J. O., Barry, P. L., Hunt, M. R., and Groth, T., *Clin. Chem.* 27, 493 (1981).

Statistical Models in Quality Control for Trace Analysis

Robert S. Elder

FSIS Science, Room 612 Annex
 U.S. Department of Agriculture
 Washington, DC 20250

1. Introduction

It is well known that there are important factors affecting accuracy in trace analysis, such as handling loss, contamination and purity of reagents.