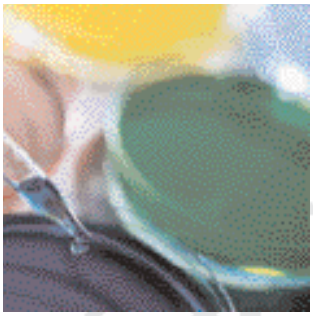
Analysis of Environmental Microbiology Data from Cleanroom Samples

David Hussong* and Russell E. Madsen



PHOTODISC, INC.

Current microbiological methods cannot measure microbial contamination at the levels that engineers and regulators seek to establish for aseptic processing cleanrooms. New approaches for assessing data and establishing alert and action levels are advocated, and an example of one analytical tool is considered.

David Hussong, PhD, is a microbiologist at FDA's Center for Drug Evaluation and Research, Office of Pharmaceutical Science, 5600 Fishers Lane, Rockville, MD, 20857, tel. 301.827.7490, fax 301.827.3084, hussong@cder.fda.gov.** **Russell E. Madsen** is president of The Williamsburg Group, LLC (Gaithersburg, MD).

*To whom all correspondence should be addressed.

**The comments contained in this article are those of the authors only and do not necessarily represent the positions or policies of the FDA.

In February of 1993, Judge Wolin of the US District Court in New Jersey issued a ruling, known as the Barr decision (1), through which the term *out-of-specification* (OOS) entered common use, even though the meaning of the term remained ambiguous to regulators and the industry. For manufacturers, this decision was rational and precedent-setting because it defined procedures and criteria used for the production and release of drug products (2). However, subtle differences in definition and changing regulatory opinion since the ruling have led to confusion and indecision when handling microbiological data. There have been many contributing factors, but much of the problem lies in interpretation by individuals lacking a background in practical microbiology. Too often, conclusions have been drawn with the intent to satisfy regulatory "requirements" that do not exist, leading to improper actions (and perhaps policy) on the basis of speculation.

In the Barr decision, the judge preferred the term OOS even though FDA was more accustomed to using the term *failure investigation*. However, the results discussed in the ruling extended beyond failure investigations of finished products to include practices related to validation. In the Court's ruling, OOS results were grouped under three events: laboratory (i.e., analytical/technical) error, operator or nonprocess-related error, and process-related manufacturing error (e.g., improper standard operating procedures). The ruling also addressed single and multiple OOS results and provided procedures to follow in response to each. Testimony concerning microbiology tests supported the appropriateness of retesting and averaging because of the nature of microbiological data. FDA attempted to provide additional guidance following the Barr decision, and drafted the "Guidance for Industry—Investigating Out-of-Specification (OOS) Test Results for Pharmaceutical Production" in September 1998 (3). Concerning microbiological data, the guidance only addressed the averaging of results as described in *USP* <81>, stating, "the USP prefers the use of averages because of the innate variability of the biological test system." Averaging is fundamental to microbiological counting methods. For example, the average number of colonies counted on multiple plates determines the plate count of the sample.

The working definition of *OOS* was narrowed when the International Conference on Harmonisation of Technical Re-

quirements for Registration of Pharmaceuticals for Human Use (ICH) defined *specification* in its Q6A guidance. Under Q6A, a specification is defined as "a list of tests, references to analytical procedures, and appropriate acceptance criteria that are numerical limits, ranges, or other criteria for the tests described. It establishes the set of criteria to which a drug substance or drug product should conform to be considered acceptable for its intended use" (4). By this definition, "OOS" does not apply to process control tests unless the test includes assaying components of the product. Historically, however, the scope of OOS has gone beyond the ICH definition and has been applied to more than just product specifications. This historical misconception has led to an inappropriate use of environmental microbiology results as surrogate release criteria. This is especially troublesome when establishing alert or action levels at very low quantitative levels.

Table I: The standard deviation (SD) of the average plate count.*

Mean count	% SD
1	100
3	58
5	45
10	32
25	20

* SD is expressed as a percent of the average count.

The *European Pharmacopoeia*, section 2.6.12, "Microbiological Examination of Non-Sterile Products (Total Viable Aerobic Count)" recognizes the inherent precision limitations of the enumeration methods and allows a test value exceeding the monograph limit (or acceptance criterion) by not more than a factor of five. Similarly, *USP* <51>, "Antimicrobial Effectiveness Tests," notes that variations in test values may exist when multiple samples are collected over time, and allows count variability in logarithmic intervals (0.5 $\log_{10}$) for selected results (5).

In the evaluation of environmental quality, a belief exists that critical zones in aseptic processing areas are improved when microbiological acceptance criteria are closer to zero. For example, it has been proposed that the quantitative limit for air samples not exceed 3 cfu/m^3 (*USP* <1116>). Yet some have argued this limit should be reduced to 1 cfu/m^3. These values offer no practical difference in individual samples, and a very large number of samples (and a large volume of air) would be required to demonstrate a difference in counts at this population density. Essentially, these cannot be determined in one test with current technologies. Most microbial count data from Class 100 environments reveal no colonies at all, and perhaps a more meaningful interpretation can be made by evaluating the portion of samples that yield growth. For example, a cluster of 5 samples with 1 colony may have more significance than a single sample that yields 5 colonies. Also, large differences exist between observing many samples with no colonies and establishing action levels of 1 cfu/test.

Quantitative data offer various challenges to the microbiologist. There are differences in the variation of microbiological data resulting from sample variation, population variation, and assay variation. In addition, these counts are subjective because of the influence of sampling, growth media, and incubation parameters. Except for the sterility test, the assessment of these data has a subjective element that frustrates clear, objective decision making. Acknowledging that subjectivity is the first hurdle for the analyst. *USP* <1227>, "Validation of Microbial Recovery from Pharmacopeial Articles," offers suggestions for the interpretation of results from various conventional tests that begin with a common and fixed sample and population (5). These tests use freshly grown laboratory strains that are uniformly suspended in a solution. Many analysts mistakenly assume that if one can distribute laboratory samples from a single dilution tube and obtain values for multiple culture plates that conform to a tight range (small value for sigma), then multiple environmental samples should also conform to that same tight range. This assumption confuses sample variation and assay variation. Test methods are often validated using cultures prepared on the same growth medium used in the method that will be tested. However, even from an ideal sample that is well dispersed, plate count data yield large variations when the average of the counts is small. For samples conforming to a Poisson distribution, Eisenhart and Wilson (6) showed the standard deviation, expressed as a percent of the average count, could be estimated according to the equation:

$$\pm\left(\frac{100}{\sqrt{T}}\right) \qquad [1]$$

in which T equals the colony count.

Using this equation for the deviation in plate counts reveals the fallacy of relying on small values for alert and action levels for cleanroom environmental samples. As shown in Table I, the standard deviation can be equal to the count. Interestingly, this equation provides the same estimates of error that are offered today in *USP* <1227>.

In practice, microorganisms from pharmaceutical samples are diverse, and are not in the same robust metabolic state as the laboratory-contrived samples. Furthermore, the variable distribution of microorganisms in these samples may produce a wide range of counts, and when those counts are small, they can vary by multiples. As a practical approach, a reproducible method is selected to ensure that the results are comparable. In the absence of comparable data, trend analysis becomes a futile exercise.

It is helpful to consider the history of environmental microbiology culture methods and the usefulness of environmental microbiology data for monitoring pharmaceutical manufacturing processes. The development of these methods is closely linked to the data produced and its analysis. Quantitative microbiology has its origins at the dawn of the twentieth century, when studies of water quality were found to be predictive hygienic indicators (7, 8). Hygienic studies relating to counting bacteria were encumbered by the temperature limitations of gelatin, which was used to solidify media, because the gelatin liquefied at the warmer temperatures used to incubate mammalian pathogens. Gelatin and potatoes were used to cultivate bacteria on solid media until agar was introduced in 1881 (11). By that time, enrichment cultures and differential tests were employed in counting methods that used broth cultures for the detection and estimation of bacteria, especially coliforms. Broth cultures were sensitive and useful for estimating the population

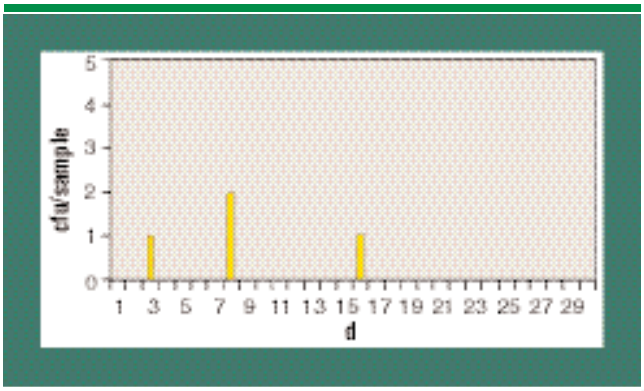


Figure 1: Colony counts for a cleanrooom site during a 30-day period.

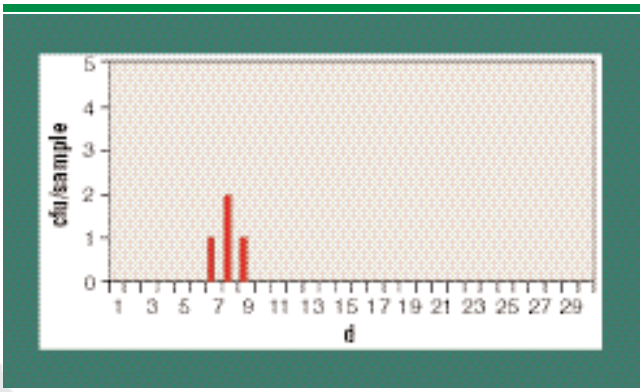


Figure 2: Colony counts for a cleanrooom site during a 30-day period.

$$\pm\left(\frac{100}{\sqrt{T}}\right)$$

of microorganisms, particularly when population densities were low. This method used dilution to extinction, and the reciprocal of the greatest dilution yielding a positive culture was the estimation of the population (12). This method emphasizes the logarithmic nature of microbiological count estimates in environmental microbiology.

Refinements to the "reciprocal of the dilution" end-point count were introduced by McCrady when he proposed that a better estimate of the count could be made by a statistical method using multiple dilutions, rather than a strict proportion (11). Subsequently, Halvorson and Ziegler demonstrated that the plate count revealed only a portion of the population in a natural sample because of sample interference, and larger estimates were made by direct microscopic observations or by dilution techniques in broth (12). Halvorson and Ziegler also proposed procedural refinements and developed what we now consider the most probable number (MPN) using multiple tube, parallel dilution-to-extinction schemes (13). This method showed consistency of the estimate regardless of the population density. Eisenhart and Wilson showed that the log of the MPN conformed to the standard symmetrical Poisson distribution (4). We have asserted that counts from cleanroom data do not conform to a symmetric Poisson curve. Furthermore, these data (zeros, ones, twos, and threes) cannot be analyzed as log values (there is no $\log_{10}$ that equals zero). Average values (not equal to zero) might be an alternative.

On the basis of these observations, the establishment of microbiological alert and action levels may require a new approach that relies less on colony count values. This is largely a result of engineering capability that has moved beyond the ability of quantitative microbiology methods for measuring and makes it necessary to evaluate quantal data or use grand averages. As a caution, statistical approaches to microbial counts in cleanrooms are difficult because these data are limited by sample volume and the results are often "zero." As shown previously, a count of 1 cfu will produce a standard deviation that is 100% of the count. This means a standard curve would include 95% confidence limits in "negative numbers" of cfu, so the one-tail statistic is needed. Furthermore, the probability of a "zero cfu" count in a non-sterile environment might approach zero, but it will never be reached. Working at these levels of counts, a single sample's result of 1, 2, or 3 cfu does not suggest any differential value. The process is "in a state of control" when a reasonable amount of data show consistent performance.

Assessment of consistent performance might include "frequency of detection," or "cumulative count averages." Microbiological monitoring results for ISO 5 and 6 (grades A and B) areas generally consist of a series of zeros interspersed with an occasional count of one or two. This condition tends to preclude the use of "cfu count" alert and action levels because virtually any excursion >1 cfu might appear abnormal. In the following example, a simple method is used to construct a control chart for data that show relatively little variation (Figure 1).

From these data, control limits might be estimated by the formulas:

$$\left.\begin{matrix}\text{UCL}\\ \text{LCL}\end{matrix}\right\} = \bar{x} \pm 3\sqrt{\bar{x}} \qquad [2]$$

in which

$$\bar{x} = \frac{4}{30} \qquad [3]$$

total cfu/number of samples, and

$$\left.\begin{matrix}\text{UCL}\\ \text{LCL}\end{matrix}\right\} = 0.1333 \pm 3\sqrt{0.1333} =$$
$$0.1333 \pm 1.0953 = \begin{cases}1.2286\\ 0\end{cases} \qquad [4]$$

This example suggests an upper control value of 1.2286, but actual data will yield only whole numbers for individual counts. The analyst could round to one or could justify the next whole number (two).

In addition to numerical value, the pattern of the data is also significant. Consider the example shown in Figure 2. In this case, the colony count values are the same as in the Figure 1 example, but they occur close together. This indicates that a control deviation may have occurred on the days with the measured counts even though the "statistic" gave the same values as the data in Figure 1. The temporal association of the counts in Figure 2 also shows a period of variation, even though the "statistic" led to the same conclusions as the data in Figure 1.

Microbiological monitoring results for ISO 7 and 8 (grades C and D) and support areas may be more amenable to the use of cfu-count-based alert and action levels becauses values other than zero occur more frequently.

It is nearly impossible to the determine assignable cause of a single microbiological excursion. This is analogous to epidemiological investigations; when a single illness event occurs, identifying the source is unlikely. Only when data are collected that relate time or location to multiple microbiological observations can a meaningful conclusion be drawn. Similarly, when a single event occurs, it should be noted and adequately catalogued to enable a comparison when other anomalies occur. This type of analysis is suited for atypical isolates (qualitative data), so rational data storage and retrieval systems are needed to enable this system to work. These were once created with notebooks and cards, but now laboratory information systems are making these records available by electronic means.

Investigations into microbiological data that exceed the environmental "trend" should be carefully considered. However, variation in microbial counts is an expected phenomenon and the only mechanism to prevent this is to require the sample to be sterile. The understanding of when and how to act with respect to microbiological results is an important skill for the pharmaceutical scientist, and is an appropriate part of written procedures that apply to microbiology results. Unfortunately, the analyst can never be absolutely certain where to place "control limits" for events which require determination of assignable cause. There remains a degree of subjectivity in this analysis because unrealistic assumptions are needed to apply a statistical analysis of microbial counts in the cleanroom environment (14).

When an abnormal environmental monitoring result (i.e., one that exceeds a control level) is encountered, its effect on product quality should be carefully evaluated. As stated above, it is unlikely that the cause of the excursion will be determined. Previous occurrences, sample type and location, proximity to exposed product and product contact surfaces, and other factors must be evaluated. Environmental monitoring data from successful media fills can serve as a reference point when evaluating the potential effect of abnormal product-associated environmental monitoring results. A decision regarding product release in these circumstances is likely to be shrouded in uncertainty.

For these reasons, environmental count results should not be used as product release criteria. It is recommended that different approaches to data analysis be explored for establishing when data deviations are not the random perturbations of a controlled system. Some old methods, including methods more than a century old, may be useful as well. Furthermore, the search for assignable causes of nonrandom deviations should be emphasized for use in identifying needs for process improvement. Within this context, the exploration of new process analytical technologies for environmental microbiology also should be aggressively pursued.

Conclusion

Current microbiological acceptance criteria for the monitoring of cleanrooms used for aseptic processing exceed the technical limits for traditional plate counting based on traditional statistical evaluation. Historically, hygienic assessments used quantal counting methods (frequency of detection) and probabilistic evaluations of the detections to predict microbial density. New counting technologies may overcome the limitations of cultivation-based counting, but new perspectives regarding data analysis should also be considered, possibly integrating the approaches from early hygienic microbiology. These new perspectives and technologies may provide a more useful assessment of process control and identify conditions that can benefit from investigation and process improvement.

References

1. *US District Court for the District of New Jersey v. Barr Laboratories, Inc.* 1993. Civil Action No. 92–1744 (available online at http://www.gmp1st.com/).
2. R.E. Madsen, "U.S. v. Barr Laboratories. A Technical Perspective," *PDA J. Pharm. Sci. Technol.* **48**, 176–179 (1994).

3. FDA, "Guidance for Industry—Investigating Out-of-Specification (OOS) Test Results for Pharmaceutical Production (draft guidance)," (FDA, Center for Drug Evaluation and Research, Rockville, MD, September 1998).
4. FDA,"International Conference on Harmonisation; Guidance on Q6A Specifications: Test Procedures and Acceptance Criteria for New Drug Substances and New Drug Products: Chemical Substances," *Federal Register* **65** (251), 83041–83063 (29 December 2000).
5. *USP 27–NF 22 (*US Pharmacopeial Convention, Rockville, MD, 2004).
6. C. Eisenhart and P.W. Wilson, "Statistical Methods and Control in Bacteriology," *Bact. Rev*. **7** (57), 137 (1943).
7. P. Miquel, *Manuel pratique d'analyse baclériologique des eaux,* (Gauthier-Villars et fils, Paris, 1891).
8. S.D. Gage, "A Study of the Numbers of Bacteria Developing at Different Temperatures and the Ratios Between Such Numbers with Reference to their Significance in the Interpretation of Water Analysis; Biological Studies by the pupils of William Thompson Sedgwick, Boston (1906)," in S.C. Prescott, C-E. A. Winslow, and M.H. McCrady, *Water Bacteriology*, (Wiley & Sons, New York, NY, 6th ed., 1946.)
9. W. Hesse, "Uber die quantitative de Bestimmung der Luft enthaltenen Mikroorganismen," *Mitt. a.d. Kaierl. Gesh. Berlin*. **2,** 182–207 (1884).
10. E.B. Phelps, "A Method for Calculating the Numbers of *B. coli* from the Results of Dilution Tests," *Am. J. Pub. Hyg.* **18** (141), 145 (1907).
11. M.H. McCrady, "The Numerical Interpretation of Fermentation-Tube Results," *J. Infect. Dis.* **17,** 183–212 (1915).
12. H.O. Halvorson and N.R. Ziegler, "Application of Statistics to Problems in Bacteriology. I. A Means of Determining Population by the Dilution to Extinction Method," *J. Bacteriol*. **25,** 101–121 (1933).
13. H.O. Halvorson and N.R. Ziegler, "Application of Statistics to Problems in Bacteriology—III. A Consideration of the Accuracy of Dilution Data Obtained Using Several Dilutions," *J. Bacteriol.* **26,** 559–567 (1933).
14. H.R. Neave and D.J. Wheeler, "Shewhart's Charts and the Probability Approach," presentation at the Ninth Annual Conference of the British Deming Association, Birmingham, England, 15 May 1996.**PT**