AOAC INTERNATIONAL
Presidential Task Force on
Best Practices for Microbiological Methodology
US FDA Contract #223-01-2464, Modification #12

Executive Summary
Sampling Working Group (SAWG)

INTRODUCTION

As with any type of testing, an understanding of the sampling and measurement procedures for microbiological methods is necessary for gaining confidence that the obtained results "represent" the intended population or fulfill a study's purpose. The confidence of results can be undermined if care is not taken to control and minimize the variation of observed results due to sampling, sample preparation and measurement. To address this concern, the AOAC has asked the Sampling Working Group (SAWG) of the BPMM Task Force to identify and address the components of sampling and measurement variation – specifically, the factors that contribute to and must be controlled or understood in order to gain an understanding of results and thereby enhance their proper use. This would include identifying components across the whole process of sampling and measurement, including the method of measurement and the laboratory performance. Once these components of variation are understood, proper application of the method can be designed.

There has been significant work done by the International Commission on Microbiological Specifications for Foods (see ICMSF, 2002) to develop and provide guidance on the use of microbiological sampling plans for foods. The statistics underlying these sampling plans, however, are not well understood (Dahms, 2004). The components of variation, referred to above, were not considered in determining the operating characteristics of the plans; instead, rather idealized assumptions were made.

In view of these issues, the objective under consideration by the SAWG of the AOAC International Best Practices for Microbiological Methods (BPMM) Task Force is: (Contract question #3) What are reasonable performance standards (criteria) when microbiological methods are to be used for: 1)Attribute testing, 2) Variables testing, and 3) Process control testing.

METHODOLOGY

The AOAC objective set forth is broad, and therefore the SAWG narrowed its scope to identify important areas that could lead to further investigation. Certain assumptions were made. One primary underlying assumption is that a statistically representative sample can be obtained and that if composite samples are to be used, then these composites will be "representative" from a unit or amalgamation of multiple units that they are to characterize. An indication that a set of samples is representative of the lot is that the variation between samples is less than the mean. It also follows from these

assumptions that outright "errors" due to mislabeling of samples, cross-contamination, incorrect readings from a machine, etc. would not be addressed. These possibilities are important to consider, and should be part of any well-designed laboratory standard operating procedure (SOP), but are beyond the scope of the SAWG.

The issues to be addressed by the group do not depend, per se, on whether the type of test being considered is an attribute or variable test. In other words, the recommendations presented below are being made with regard to qualitative tests as well quantitative tests that are more familiar to AOAC. In lieu of the above discussion, the SAWG considered the following tasks:

1) Identify and address performance components of variation relative to intra-laboratory, and inter-laboratory performance.
2) Identify and address components of variation of measurement error associated with the method within the laboratory.
3) Identify process control statistics and recommend a set of performance standards for statistical process control using microbiological measurements.

I. Components of variation relative to intra-laboratory and inter-laboratory performance.

The SAWG believes that to determine method performance, controlled inter-laboratory studies are needed. The recommendations are closely aligned with AOAC recommendations for collaborative studies of chemical analytical methods. The recommended performance standards are:

1. Ruggedness tests should be performed that attest to the robustness of the analytical procedure under expected normal operating procedures. Ideally 5-7 critical steps of the procedure should be identified, and the nominal, upper and lower specs for each step evaluated.

2. Microbial test validation should include estimates of test sensitivity, specificity, and accuracy.

3. A Collaborative study consisting of 5-10 laboratories should be conducted to determine reproducibility and repeatability standard deviation measures that cover the range of levels expected to be encountered and that are of regulatory interest. If this is not possible, then at least an intra-laboratory study, using more than one analyst, separated from each other, should be conducted. From these results, formulas predicting the standard deviations as a function of level should be estimated.

4. For QA purposes, laboratories should establish a range of acceptable results for individual samples based on confidence intervals using the repeatability standard deviations. Also, laboratories should establish process control

procedures, and use statistical process control methods for tracking performance over time.

5. When reporting results, the range given as the 95 percent confidence interval on the measurement should be stated.

II. Identify components of variability within the lab.

The SAWG focused on examples of method protocols to examine where the measurement error variation can occur. Enclosure A presents a detailed account of our identification of major sources of sampling variation that occur within the laboratory. We are recommending that laboratories develop a protocol for maintaining process control at critical points of the analytical procedures. The recommended performance standards for laboratories are:

1. Establish a process for listing sources that contribute to the variability of results in the laboratory (this should be developed).

2. Perform intra-lab repeatability studies to determine statistical distribution of results associated with the sources of variability.

3. Establish statistical process control procedures (based on split or check samples) within the laboratory to monitor performance.

4. For methods that involve confirmation of particular types of organisms where interfering organisms are expected, conduct a study to determine the proportions of targeted and interfering organisms in samples. This will help determine how many confirmations are needed to minimize false negative outcomes.

III. Statistical Process Control (SPC).

SPC is a very broad area which SAWG believes is not well known to the scientific community. Consequently, for this task, the SAWG presents a general introductory discussion (Enclosure B) together with numerous examples. The suggested performance standards are general principles that should be followed, representing normative practice. These are:

1. Charts of plots of the output data are necessary for gaining the full benefit of doing SPC.

2. When the process is under control, the results plotted on a statistical process control chart should be normal or nearly normally distributed. In cases where this is not true and an alternative known distribution cannot be determined, transformations of the data should be considered.

3. During some "initial" period of time, when it is presumed the process is operating in a relatively stable manner – or is in control, the distribution of the measurements should be estimated and rules for evaluating the process should be formulated. Use of about 20-30 results (samples) or more for computing means and standard deviations or other summary statistics needed for distribution estimation is a desirable goal. However, this stipulation can be relaxed and thus should not hinder or limit the use of control charts if resources do not permit, in a timely fashion, analyzing this number of samples.

4. Rules for evaluating process control should be set with aids assessing the two types of errors: Type I (α-probability), declaring the process out of control when it is not, and Type II (β- probability), not declaring a process out of control when it is. Typically there are two measures that are used for assessing these errors: 1) the probabilities of the two types of errors at a given time and 2) the average run length (ARL) or expected number of samples before an out of control signal (one of the rules being not met) is seen. When developing rules, the α-probability (Type I error) should be kept low, for example, below 1%, or the ARL should exceed 100 (corresponding to less than 1% α - error).

5. When a process is thought to be "in control," the limits for assessing individual results are set at a distance from the mean (target), expressed as standard deviation units from the mean or process target value. The recommended and default distance is 3 standard deviations. Additionally, characteristics related to food safety may be targeted more than three standard deviations above or below critical limits, however statistical process control limits should still be placed 3 standard deviations from the target value.

6. There are numerous run/ trend rules that can be used, such as runs test, moving averages and CUSUMS, for detecting shifts in the process mean; and rules for detecting shifts in the process variation or other auto-correlated patterns that could be due to a systematic source of variation. The use of any of these may depend upon particular expected conditions when the process is out of control.

7. Specification Limits are not Statistical Process Control limits. Specifications are either customer, engineering, or regulatory related. Specification limits should not be placed on a control chart insofar as these might be considered as process goals thus influencing the efficacy of SPC procedures for ensuring a controlled process, and thereby undermining the safety of the product.

For more details concerning the specific performance criteria, please review the referenced Enclosure materials.