

MQOs – Straw Proposal with Comments

Prior to the MQO subgroup call on 2/14/06, Nan Thomey and David Kimbrough sent comments on the EPA MQO Straw Proposal. John Phillips sent comments after the call. Their comments have been inserted into EPA's original document for your information, with their names in advance of their comments. If viewed electronically, EPA's straw proposal is black/white; Ms. Thomey's comments are highlighted in yellow; and Mr. Kimbrough's are highlighted in aqua.

David Kimbrough: General note. In a routine CWA compliance sample, it is very rare for a laboratory analyst to know if a positive result is a false positive or a negative result is a false negative. This is usually determined not by examining the sample itself, but by examining the quality control samples in the same batch. False positives are assessed by measuring unspiked blank samples (distilled water usually). False negatives are assessed by spiked blanks (LFB) and spiked matrix samples (LFM). It is assumed that if there is a problem with the blank samples, similar problems will occur in the compliance samples. The reverse is not assumed however. So when we talk about false positives and negatives, we're talking about blank samples as surrogates for real CWA compliance samples.

1. Alpha, α

1. **Definition:** The tolerated probability of a "false positive" (i.e. Type I error).
False positive -- Concluding that the analyte is present when in fact it is absent.

David Kimbrough: This definition is problematic. We never know if an analyte is "absent". In fact, most 40 CFR 136 analytes can be found in just about every CWA compliance sample. They are generally at concentrations too low to be measured. From an operational perspective there is no way to use this definition. An operational definition for a false positive is more accurately defined as a result above the reporting limit (Ld or Lc) when in fact it is below.

2. **Straw Proposal:** α goal = 1% (John Phillips) at the Lc. It is generally well accepted that this can be achieved.

Nan Thomey: Labs seem to be in agreement with this. Comments have included observations that this may change the statistics that are applied to the standard deviation used in some of the current procedures to achieve this objective, but that should not be problematic. Also, a comment was made that if we take this approach, we probably do not want to continue the Lc/Lq denotations since it would be contradictory. Suggestions included using the terms MDL and ML or MQL if we change this criterion.

David Kimbrough: False positives are measured in unspiked blanks. Why should we accept any false positives in an unspiked blank? The easy majority of 40 CFR 136 methods do not produce non-zero values when an unspiked blank is analyzed (unless of course they are "accidentally" spiked, i.e. contaminated). Most of the remaining uncensored methods can just as easily produce a negative or zero value as a positive value. Given all of this, why would we accept any false positives in this situation?

2. Beta, β

1. **Definition:** The tolerated probability of a “false negative” (i.e. Type II error).
False negative -- Concluding that the analyte is absent when in fact it is present.

David Kimbrough: Once more the absent / present paradigm is a problem. Laboratories almost never report results as “absent”. They report them as <less than or ND. Our definition should match what is actually done. An operational definition for a false positive is more accurately defined as a result below the reporting limit (Ld or Lc) when in fact it is above.

2. **Straw Proposal:** β goal should not be established at this time. This is likely to fall out of the procedure that is finally selected by the FAC. To choose a goal now would prejudice the FAC’s choice of the final procedure. For example:

1. β goal of MDL is 50% (**John Phillips**) at the L_c
2. β goal of MRL is 5% (**John Phillips**) at the L_c
3. β goal of IDE/IQI is 1% (**John Phillips**) at the L_c
4. β goal of Hubaux-Vos is selected (**John Phillips**) at the L_c

Nan Thomey: A goal for β for the MDL (L_c) should definitely not be set. Although in theory β at L_c is 50% when α is 1%, this is only when many assumptions are met (constant variance, 100% recovery, no interfering qualitative identification criteria, etc, etc). These assumptions are not met in practice so β does not = 50% when $\alpha = 1\%$. The alpha criterion is much more important, so β should be ignored for L_c (MDL).

David Kimbrough: False negatives are measured in spiked blanks. Why should we accept any false negatives in a spiked blank? We know that the analyte is present, we put it there. I see no reason to accept any false negative under these conditions. The only caveat I would add is that if the spiked blank is exactly at the censoring limit (L_c or L_d) then the results should not be censored for the spiked blank.

3. Accuracy

1. **Definition:** The degree of agreement between an observed value and an accepted reference value. Accuracy includes a combination of random error (precision) and systematic error (bias) components, which are due to sampling and analytical operations; a data quality indicator.
2. **Straw Proposal:** Accuracy goal is $\pm 20\%$. This is a goal for validating methods and can be tightened or widened depending on the analytical method. For example, the development of QC for GC/MS method 525 indicated the accuracy window should be $\pm 30\%$ to accommodate the disparate range of organic compounds (pesticides, Aroclors, etc.) covered by this method. After a method is validated for an analyte, either the CFR or the listed compliance method will specify an accuracy “band” that must be met by laboratories.

Nan Thomey: I think the labs want to discuss the accuracy criterion further before deciding. It is more complicated due to differences in what are

realistically achievable for different methods and/or analytes. The choice of how we will demonstrate this criterion will also be an important consideration for laboratory operations.

David Kimbrough: I think we can tolerate much more bias at L_q provided L_q is significantly higher or lower than the decision level.

Case 1: $L_c = 0.1$, $L_q = 1$ ppb, and $WQBEL = 10$ ppb Thallium. If we have a sample with 2 ppb thallium, suppose we actually read it as either 2.8 or 1.2 ppb (+/- 40% bias). If you round to one significant figure, these results are all the same. Even if you don't, that level of bias is acceptable since it does not change the regulatory decision making: it is still a detect but not a violation. So although 40% sounds like a lot in a relative sense, in the context of regulatory decision making, it is not.

Case 2: $L_c = 0.1$, $L_q = 1$ ppb and $WQBEL = 0.01$ ppb for Cyanide. If we have a sample with 2 ppb CN^- , suppose we actually read 2.8 or 1.2 ppb (+/40% bias). Again it does not matter much since either reading is clearly a violation.

Case 3: $L_q = 2$ and $WQBEL = 2$ Nitrate. Again, if the sample is 2 ppb, and the results are read as either 2.8 or 1.2, it makes a huge difference, the difference between compliance and violation.

John Phillips: Accuracy goal is $\pm 10\%$ at the L_Q .* This value may be adjusted up or down depending upon the regulatory or non-regulatory use. One use may be the validation of methods, when a method is proposed for use in the CWA program.

4. Precision

1. **Definition:** The consistency of measurement values quantified by measures of dispersion such as the sample standard deviation. Precision must be defined in context – e.g., for a certain analyte, matrix, method, perhaps concentration, lab or group of labs.
2. **Straw Proposal:** Precision goal is $\pm 20\%$. This is a goal for validating methods and can be tightened or widened depending on the analytical method. Again for EPA Method 525 this window is $\pm 30\%$, and after validation for an analyte, either the CFR or the listed compliance method will specify a precision “band” that must be met by laboratories.

While we can set goals for initially deciding whether to thoroughly develop and validate a specific analytical method, these goals are not absolutes. Ultimately the final DQOs/MQOs are tailored to the intended use of the analytical method, and a best estimate on how well good laboratories can be expected to perform with that method, analyte and matrix combination.

3.

Nan Thomey: Assuming this means 20%RSD, it is a good goal and may also carry the benefit that L_d is approximately equal to L_q at that level of precision.

David Kimbrough: I not convinced we really need to measure precision per se at L_q . If we get good accuracy on a batch-by-batch basis, I do not know that long term precision will tell us anything that we do not already know.

John Phillips: Precision goal is $\pm 10\%$ at the L_Q .* This value may be adjusted up or down depending upon the regulatory or non-regulatory use. One use may be the validation of methods, when a method is proposed for use in the CWA program.