Measurement Uncertainty: Maxims by Colle (c. 1998)

1. BEWARE OF RED-FLAG WORDS like "error", "systematic", "bias", "precision", etc. There may be some chance that these words are being used correctly, but chances are much greater that they are not. If you see them in a guidance document, there is an even greater probability that they violate accepted international practices.

2. DON'T TALK TO STATISTICIANS ABOUT METROLOGY OR UNCERTAINTIES. Firstly, few, if any, statisticians understand metrology and none that I am aware of understand the assessment of measurement uncertainties. The assessment of measurement uncertainties is not a statistical issue and should never be treated as one. Consult statisticians to help you design parts of experiments, or do statistical interpretations of data, but don't ever trust them to design a measurement uncertainty treatment model for you.

3. MAKE YOUR MEASUREMENT MODEL EXPLICIT. If you made a measurement (or have a prescribed measurement method) then you HAVE a measurement model ~~when~~ whether you realize it or not. You must understand it fully to do an adequate uncertainty assessment. Once your measurement is done, your uncertainty is set and determined. The trick, of course, is to determine what this uncertainty is. This is what the understanding of your model is all about. And if you think about your model before you make the measurement well then you can actually plan your experiment, optimize the conditions, and figure out ahead of time as to how you are going to assess the uncertainty components.

4. THERE ARE NO UNCERTAINTY TYPES. There are only different types of methods used to evaluate the uncertainty components. If you find yourself still thinking in terms of "random" things and "systematic" things than you are doing something wrong and you don't understand what uncertainty assessments are all about.

5. DON'T EVER USE THE SQUARE ROOT OF THE NUMBER OF COUNTS AS AN UNCERTAINTY ESTIMATOR. If you find yourself doing this, then you are probably doing something pretty dumb. Informing statistically-illiterate physical scientists that the radioactive decay process is controlled by the Poisson distribution is one of the largest disservices ever imposed on our discipline. The MEASUREMENT of radioactivity is almost never, ever Poisson distributed. All books and chapters containing all of the different ways (and conditions) to make the Poisson approximations should be left as cute student exercises and should not be used by any serious metrologist. These sources (like the MARLAP manual being prepared) rarely, if ever, spend any time on telling you all of the times when the approximations are not valid. It should be pretty obvious that the square root of N tells you nothing about the measurement process and the understanding of that process is largely what an uncertainty assessment is all about.

6. USE CANONICAL VALUES FOR UNCERTAINTY ESTIMATES WHENEVER POSSIBLE. It doesn't make sense, but people all the time have oodles and oodles of data on the variation of their measurement process, yet they ignore it, and instead will base their uncertainty on a calculated standard deviation for a few measurements most recently taken. This doesn't make sense but is done all the time.

7. YOU MUST CLEARLY UNDERSTAND (&YOUR UNCERTAINTY ASSESSMENT MUST REFLECT) EXACTLY WHAT YOUR CONDITIONS OF REPLICATION MEAN. The extent to which conditions for a type A assessment are allowed to vary over successive "repetitions" of the measurement process determines the scope of the statistical inferences that may be drawn from the measurements (and their reported uncertainties).

8. DON'T USE THE STANDARD DEVIATION OF THE MEAN AS AN UNCERTAINTY ESTIMATOR UNLESS YOU HAVE VERIFIED THAT ITS USE IS APPROPRIATE. This is one of those few times when a statistician might help you. In general, when multiple measurements of a particular quantity made on a single occasion exhibit closer mutual agreement than measurements made on different occasions so that differences between occasions are indicated, the value of a computed standard deviation of the mean obtained by lumping all of the measurements together will underestimate the actual standard deviation of the mean. Essentially every measurement method used by this laboratory that I have seen over the past 25 years exhibits a statistically-significant within- and between-measurement-occasion (or within- and between- sample) difference in variability. Yet, the blind practice of using the standard deviation of the mean is rampant. Nested uncertainty treatments are widely available in the literature.

9. UNCERTAINTY ESTIMATES OF ANY KIND ARE VERY UNCERTAIN. That is to say, that the uncertainty on an uncertainty is usually huge. This is equally true for both Type A and type B evaluations. One must remember that computed statistics like standard deviations with few degrees of freedom are being sampled from very broad distributions. It is generally safe to assume that any measurement uncertainty is itself uncertain to at least 50 % or more. Standard deviations for typical radioactivity counting experiments (with negligible "counting statistic" uncertainties) for $n = 5$ replicates often vary by a factor of two from trial to trial. This is where the evaluation of canonical values for the measurement process becomes useful.

I SHALL UPDATE THIS LIST OF MAXIMS TO CONSIDER AS THEY OCCUR TO ME.