April 2008
What Have We Learned in the Last Two Decades?

Cleaning validation for pharmaceutical manufacturing is no longer in its infancy. Several months ago in a Cleaning Memo I discussed two papers on cleaning validation that were published in the 1980's. Different dates could proposed for the official "birthday" of cleaning validation, but the early 1990's are fairly close, with the Barr Labs decision and the FDA's cleaning validation guidance document. One would think that we would have learned something over this time to make cleaning validation more effective and more efficient. Below is a list of things (by no means exhaustive) of possible areas where we can spend less time going forward so that we can focus on those items of more concern for the quality and safety of pharmaceutical products. I believe these things are consistent with regulatory bodies' risk-based approach. I might also note that these are also consistent with a statement from the 1993 FDA cleaning validation guidance, that answers to questions about the cleaning process (that is, understanding our cleaning processes) "may also identify steps that can be eliminated for more effective measures and result in resource savings for the company". Now don't think that the FDA is concerned about making you company more profitable; however, it is a reasonable expectation that they are concerned about the costs of drug products to consumers (and it is in this sense that they would like the industry to be more efficient).

Measuring bioburden in protocols
Certainly bioburden control is important. However, if I have repeatedly performed cleaning validation protocols on different products using the same cleaning process, and then a new product comes along which is cleaned by the same process, is it really necessary to sample for bioburden on the new product if my previous protocols have demonstrated consistently acceptable bioburden control? For example, if I am in bulk biotech (let's say E. coli fermentation) and have consistently demonstrated bioburden control with my cleaning process, is it necessary to perform bioburden testing on a new protein made on the same equipment and cleaned by the same process? In other words, this is like utilizing a grouping strategy, but only doing the grouping for some test parameter, namely bioburden. Of course, it is necessary to understand my cleaning process, and there may be factors that would prevent me from using this approach. For example, if I were making finished drugs and the new drug product had components that had unusual bioburden properties; I might not utilize this strategy. A concrete example might be where previously my product was from synthesized raw materials, but a new product had at least one component of natural origin.

Setting rinse sampling limits
For rinse water sampling using WFI, and particularly for final process rinse water sampling in a CIP process, is it really necessary to require that the final rinse meet the WFI specs of no more than 10 CFU per 100 mL? That specification is designed for water in the recirculating WFI loop. Once I take it out of the WFI loop and pass it through clean equipment, do I really expect it to still meet the WFI specs for bioburden? Furthermore, if it exceeds the WFI specs, even by a factor of 10, once the system drains and dries won't the normal dehydration of vegetative organisms further decrease the bioburden left on surfaces. Additionally, if I am in aseptic processing and the equipment is either steamed or parts are autoclaved, my concern is further minimized. Certainly I don't want gross bioburden (in part because of endotoxin concerns from gram negatives), but setting rinse water limits at WFI specs is generally not reasonable. [Note that it is possible to extend this argument to meeting conductivity specs for rinse water sampling for either WFI or Purified Water.]

Measuring cleaning agents in protocols
This is similar to the issue of measuring bioburden. If I am using the same cleaning process (including agent) for several manufactured products, and I have shown consistent and acceptable result for several products, if I perform cleaning validation for a new product, can I appeal to the previous data to cover the new situation? I believe, based on the understanding of the cleaning process, this could be an acceptable approach.

Performing "Clean Hold Studies" on dry equipment
The issue here is that clean hold studies are designed to measure potential recontamination of equipment on storage. There are two major concerns – one is bioburden proliferation because the equipment is stored wet. The second is external recontamination (from dust and dirt) because the equipment is not sealed, or otherwise protected. My question is, if the equipment is appropriately dry at the end of the cleaning process, and if the bioburden is acceptable at the end of cleaning, is it reasonable to conclude that bioburden proliferation is not likely to occur (or will not occur) unless the equipment becomes wet through external sources of water or through condensation because the equipment is sealed before it is cooled? Under those conditions, my clean hold time study might just involve holding the equipment for a defined time and then visually inspecting it for visual cleanness. That certainly simplifies the clean hold study, but it is a risk-based approach. In other words, I should spend more effort in assuring the equipment is dry, as opposed to demonstrating that bioburden does not proliferate on storage.

Extent of sampling recovery studies
I am of the opinion that many companies spend too much time in sampling recovery studies, such as swab sampling studies. While it may be appropriate in analytical method validation to perform five or six levels over the proposed linear range, is it required that I also spike coupons at five or six levels to determine recoveries? My answer is "No". For measuring different concentrations of solutions, I can expect the results to be linear (at least over a certain range). However, with all the variability of swab sampling, is it reasonable to expect some kind of linearity with the results? Therefore, why do five or six spiking levels? My preference is to do one spiking level, at the residue limit. In other words, if my residue limit is X μg/cm2, I would prefer to spike at a level of X μg/cm2. Based on the assumption that within the range of typical residue levels the percent recovery drops with increasing load (because of loading onto the swab), the recovery at that level (X μg/cm2) should apply to all spiked levels below that level. If the concern is variability of the swabbing procedure, then it may be appropriate to spike at one additional level, such as 20% of the residue limit. From a practical perspective, I would then choose the lower of the two recovery percentages as my "official" recovery value. Note that while I assumed that recovery is higher at lower levels, this does not mean that every time I perform recovery studies at two spiked levels, the higher spiked level will always give lower percent recovery results. It is entirely possible that measured recovery values will be lower for the lower spiked levels. This is most likely due to just the variability of the swab sampling procedure. In other words, on day one I get 68% recovery at a spiked level of X μg/cm2 and 63% recovery at a level of 0.5X μg/cm2; on day two, with the same swab operator, I get 65% recovery at a spiked level of X μg/cm2 and 71% recovery at a level of 0.5X μg/cm2. In any case, I would prefer to have more six replicates at one level (or six replicates on each of two spiked levels) than to have three replicates at each of five or six spiked levels.

This list of topics where it is possible to simplify cleaning validation based on process knowledge is by no means exhaustive. The examples given are just here as illustrations of what can be done based on better knowledge and past data.