July 2001
Product Grouping Strategies for Cleaning Validation

The purpose of "product grouping" for cleaning validation is to simplify validation in cases where there are a large number of "similar" products made on the same equipment and cleaned using the same cleaning SOP. Grouping strategies are not mentioned in the FDA's cleaning validation guidance document; however, in a "CGMP Note" the FDA has indicated they handle this on a case-by-case basis. The PIC/S validation document does specifically mention the grouping concept, although they call it "bracketing".

The principle behind grouping is to select a representative product from among a group of similar products, perform three PQ cleaning runs on that representative product, and claim that the data also support effective cleaning for other members of the same group. Needless to say, the acceptability of the concept of grouping must be mentioned in the company's cleaning validation master plan or cleaning validation policy statement. Then, the specific criteria and practices used for grouping should be detailed in a "grouping" SOP.

The first criterion for grouping is that the products must be similar products. The easiest case to deal with involves making similar drug products with the same excipients, but with different levels of actives, or with the same levels of excipients and actives except for different flavors or colorants. However, grouping is not limited to such cases. Groups can be formed by drug products with different actives and different excipients, provided that the products are similar product types (liquids, creams, ointments, gels, powders, tablets, etc.) dosed by the same route of administration (oral, parenteral, topical, etc.). The basis for such restriction is that it allows for a more consistent determination of a worst-case product (discussed later).

The second criterion is that the various products be made in the same equipment or equipment train. One cannot validate two products if one product is made in a ribbon blender and the other is made in a cone blender. All the products have to pass through the same equipment or same equipment train. The rationale for this restriction again relates to the selection of the representative (worst-case) product. Unless all products are made in the same equipment, it becomes difficult to select the worst-case product. The case sometimes arises where in one equipment train, not every manufactured product contacts every piece of that equipment train. A problem arises if the worst-case product is a product that doesn't contact every piece of equipment in an equipment train. How does one handle the cleaning validation of that one excluded item? There are several options, but it definitely complicates the grouping strategy.

A third criterion is that the products in the group all be cleaned with the same cleaning SOP. This includes the identical cleaning agent, cleaning agent concentration, times, temperatures, and other process parameters. Again, the rationale for this is to allow selection of the worst-case product.

Once a group is formed, the next task is to pick the representative product (three PQ cleaning runs will be performed on the representative product). The general principle behind selecting that representative product is that it should be the worst case for cleaning. In certain cases, the selection of the worst case is straightforward. For example, if one is dealing with similar products with different levels of the same active, then the highest active level is chosen as the worst case. If the drug products are significantly different in formulation, the principle is to choose the most difficult to clean with the proposed cleaning process. This selection process may result in a different worst-case products depending on the cleaning process (one can easily imagine that

certain products would be more difficult to clean if an aqueous alkaline cleaning agent were used, and that within the same group, other products would be more difficult to clean if an aqueous acidic cleaning agent were used). This has been handled in different ways. A common approach is to select the product with the active that is least soluble. While this approach has some merit, it may not stand up to good scientific scrutiny. For example, it is well known that in many cases of cleaning finished drug products, it is actually the excipients (rather then the actives) that are more difficult to clean. Furthermore, the difficulty of cleaning the active will vary depending on whether the cleaning process involves an alkaline, acidic or neutral pH. If one bases the decision on the solubility of the different actives, it may make more sense to base it on the solubility of the actives at the cleaning pH (and perhaps the cleaning temperature also).

The preferred method of selecting the worst-case product is to have a laboratory study comparing the cleaning performance on the different products using a simulated cleaning process. This typically involves varying things like time or cleaning agent concentration. For example, under conditions of the same temperature and the same cleaning agent concentration, the cleaned product that requires the longest to reach a defined criterion of "clean" is the worst case. Alternatively, if temperature and time are held constant, then the product that requires the longest time to clean is the worst case. Needless to say, these laboratory evaluation conditions should simulate as much as possible the actual cleaning conditions to be used in the manufacturing equipment. For example, determining the worst-case product at 40°C probably isn't relevant if the cleaning process utilizes a temperature of 70°C. Suppliers of cleaning agents may be able to assist pharmaceutical manufacturers in completing such studies.

Once the worst-case product is selected, the next consideration is to determine the appropriate residue limit for the three PQ runs using that product. A problem arises because sometimes the worst-case (most-difficult-to-clean) product has a residue limit that is above the limit for another product in the group. For example, if the worst-case product contains Active A, the calculated acceptance residue limit for cleaning of that product might be 12 µg per swab of Active A. At the same time, another product in the same group (a product which is easier to clean) might contain Active B, and the calculated acceptance residue limit for that product is 3.5 µg per swab of Active B. The assumption behind a grouping strategy is that, if one is able to demonstrate cleaning of the worst-case product down to a limit of 12 µg of Active A, then one should be able to clean all the other products in the same group (which are easier to clean) down to that same limit. However, having confidence that the second product can be cleaned down to a level of 12 µg of Active B is not enough, since for the second product one really needs to clean down to 3.5 µg of Active B.

There are two ways to handle this. One is to perform the three PQ runs on the worst case product, but rather than setting the acceptance limit at 12 µg of Active A per swab, the acceptance criteria is set at 3.5 µg of Active A per swab. The rationale for this approach is that, if one can clean the worst case product down to that level (3.5 µg per swab), then one can reasonably conclude that one should be able to clean other products, which are easier to clean, down to that same level. The alternative that is often used is to perform three PQ runs on the most difficult to clean product and three PQ runs on the "most toxic" product (that is, the product with the lowest residue limit). There are assumptions behind each strategy, but either could be used.

As a general rule, product grouping strategies are not used for newer drugs requiring an NDA. The rationale for this is that companies want their cleaning validation to be as "clean" as possible for any new product. Grouping strategies will be more used for older products in order to bring a facility under compliance as far as cleaning validation is concerned. Some companies will choose, as time allows, to perform an individual cleaning validation on each individual product in a group following the initial validation of the group.

However, in approaching a grouping strategy for cleaning validation, one should try to keep the strategy as simple as possible. If it takes significant work to justify the worst-case product, or if it unnecessarily complicates the manufacturing process by meeting all the criteria required for grouping strategies, it may be more cost effective to validate products separately.

This discussion is meant to present the basic framework and rationale for product grouping strategies for cleaning validation. It does not address all issues, such as how to handle cleaning validation when a new product is being considered for addition to the validated group.