DRUG DEVELOPMENT AND INDUSTRIAL PHARMACY, 21(11), 1257-1283 (1995)

RESEARCH PAPERS

BLEND UNIFORMITY AND UNIT DOSE SAMPLING

J. Berman[1] and J.A. Planchard[2]

Marion Merrell Dow Inc.
[1] Formulation Development Department - Kansas City, MO
[2] Production Support Department - Cincinnati, OH

ABSTRACT

Process validation for solid dosage forms must include a component that demonstrates content uniformity of the final blend. As a result of a recent court ruling this aspect of validation has received increased attention from both industry and government. In particular, the court ruled that the appropriate sample size for content uniformity testing of the final blend in validation batches is three times the run weight of the finished product. Furthermore, recent FDA communications suggest that the uniformity of the final blend should be held to a higher standard than that of the tablet in order to provide reasonable assurance that the finished product will exhibit acceptable uniformity. The purpose of this article is to communicate some problems that our firm encountered during the validation of a lower strength of a currently marketed tablet. The validated process for the marketed product has a long history of providing tablets that exhibit acceptable content uniformity at the higher strength. Extensive data from three validation batches manufactured at each of two production sites demonstrate consistently and significantly lower drug content in the lubricated granulation than in the finished product. The variability of the granulation data as characterized by the standard deviation is, for the most part, acceptable and comparable to that of the tablets. Unfortunately, these batches fail validation when subjected to the acceptance criteria suggested in recent FDA communications. We attribute the lower assay values of the granulation to sampling bias that occurs when small volumes of powder are extracted from a large V-blender with a sampling thief. In order to circumvent this problem and avoid rejecting acceptable processes we recommend an alternative acceptance criteria in which the mean assay of the powder samples is not relevant. This criteria relies solely on the standard deviation of the blend.

INTRODUCTION

Under the Federal Food, Drug and Cosmetic Act a drug is considered adulterated if it is not produced in conformance to Current Good Manufacturing Practices (CGMP's). CGMP's are defined in broad terms in 21 Code of Federal Regulations (CFR) Parts 210 and 211. The validation of manufacturing processes for pharmaceutical products is one requirement of these regulations. A properly validated process provides a high degree of assurance that the resulting product consistently meets predetermined specifications and quality characteristics. Thus, process validation is not only a legal requirement it is also a good business practice. There is little debate regarding the importance of process validation for pharmaceutical products. Unfortunately, there is less agreement concerning the specific details of the validation process.

Common sense dictates that the process validation program for a compressed tablet should include a component that focuses on the final blending step. During this step, various excipients are blended with the granulation in order to facilitate compaction and subsequent dissolution of the tablet. In some products these excipients can represent an appreciable portion of the final dosage form. It is therefore critically important to produce a uniform final blend in order to provide enhanced assurance that the finished product will exhibit acceptable content uniformity. Uniformity of the final blend, however, does not guarantee uniformity of the drug substance in the compressed tablets. Subsequent handling of the blended granulation such as discharge from the blender into drums and tablet press hoppers provide ample opportunity for particle segregation. This can lead to poor drug content uniformity in the finished product. Thus, a credible process validation program must demonstrate acceptable content uniformity of the final blend and finished product.

Process validation programs throughout the pharmaceutical industry have been influenced by the opinions recently rendered in *United States vs. Barr Laboratories* **[1].** In his precedent setting ruling, Judge Alfred Wolin defined some of the CGMP requirements for process validation of oral solid dosage forms in greater detail than is specified in 21 CFR Part 211. In particular, Judge Wolin ruled that the appropriate sample size for content uniformity testing of the final blend in validation and ordinary production batches is three times the run weight of the finished product. Based on the testimony of expert witnesses the Court felt

that the three times sample size adequately addresses the difficulties associated with sampling small quantities from large volume blends while accommodating the need for retesting. The concern is that larger sample sizes could mask inhomogeneity of the blend. Furthermore, Judge Wolin ruled that material can be sampled from either a blender or a drum as long as the manufacturer can demonstrate that the samples are representative of all portions and concentrations of the final blend.

The Court, in *United States vs. Barr Laboratories*, did not specify the criteria that should be used to evaluate the uniformity of blended granulation. Recent FDA communications [2] suggest, however, that USP Uniformity of Dosage Unit Criteria [85% to 115% of label claim and relative standard deviations (RSD) that are less than or equal to 6%] are too broad to be applied to blend validation. This is because a freely flowing powder may segregate when discharged from a blender and/or subjected to normal vibration in the hopper of a tablet press. In other words, the uniformity of the final blend should be held to a higher standard than that of the tablet in order to provide reasonable assurance that the finished product will exhibit acceptable uniformity.

Conceptually, sampling granulation from a blender or container to demonstrate content uniformity is relatively straight forward. In theory, a sample thief designed to extract small volumes of powder can be used to collect samples from a blender and/or drum. The sampling locations must be carefully chosen to provide a representative cross-section of the granulation. These locations should include areas that have the greatest potential to be non-uniform such as near the discharge valve in a ribbon blender or the trunnion region of a V-blender [3]. The resulting samples are then assayed using the same methods used to analyze the finished product. Content uniformity is established if the drug content of the samples conform to predetermined criteria.

Although simple in concept, demonstrating content uniformity of unit dose samples of powder blends is complicated by the potential for sampling bias. This bias can occur when small volume samples are extracted with a thief from relatively large volume populations. A sampling thief consists of two concentric tubes. The inner tube is solid except for one or more chambers that allow for sample collection. The outer tube is hollow and contains openings that can align with the chambers on the inner tube; it also has a sharp end to facilitate insertion

into the bulk powder. A handle, located at the top of the device, is used to rotate the inner tube within the outer tube in order to open or close the thief. Ideally, during sampling the closed thief is inserted vertically into the desired location within a powder blend. The thief is then opened; this allows the sample to flow into the sampling chamber(s) of the inner tube. The thief is then closed and the sample is withdrawn and collected.

A thief is far from an ideal sampling device [4 - 6]. As it is inserted into a powder blend it can carry material from the upper layers of the mixture downward towards the lower layers. If the blend has a wide particle size distribution percolation of fines through the coarser material can result in samples that are not representative of the bulk. The forces necessary to insert a long thief through a large volume population can be appreciable; this can lead to compaction and particle attrition. The static pressure of the bulk powder, which forces material into the sample chamber, is significantly greater at the bottom of a large container than in the middle or near the top. If the thief is not used in a perfectly vertical position the angle that it makes with the horizontal can affect the dynamics of the material flowing into the chamber. Special care must be taken to control the orientation of a non-vertical thief since the chamber may be exposed on the top or bottom surface of the device or somewhere in between during sampling. This problem is of particular concern when sampling from different locations within a V-blender where it is difficult to consistently use a thief in a vertical position. Furthermore, since a thief is a static sampling device it violates the two "Golden Rules of Sampling": (i) sample a moving powder and (ii) it is better to sample the entire stream of a flowing powder for short periods of time than a portion of the stream for the whole time [4].

All of these factors can result in product adulteration, particle attrition, segregation and overall sampling bias. Sampling bias is of particular concern during validation of pharmaceutical manufacturing processes where minute volumes are sampled from huge populations and then held to very high standards. The problems associated with conducting blend validation with small sample volumes have been discussed in the literature [7]. During process validation it is important to be able to distinguish between a non-uniform blend and biased samples from a homogeneous population. The purpose of this article is to

communicate some problems that our firm encountered during validation of the final blending step in a tablet manufacturing process.

METHODS

A program was developed to validate the manufacturing process for a product that contains either X mg or 2X mg of drug substance (DS) per tablet. These tablets are produced by compressing the same lubricated granulation used in a currently marketed product to proportionately smaller running weights. The validated process for this marketed product has a long history of providing tablets that exhibit excellent uniformity of DS at 4X mg per dosage unit. A critical component of this new validation effort was to demonstrate content uniformity of the final blend at weights that correspond to X mg of DS per sample. Obviously, uniformity at X mg of DS insures uniformity of the lubricated granulation at all higher strengths.

Three successive full-scale batches of lubricated granulation were manufactured, sampled and analyzed according to an approved validation protocol. Each of these batches were compressed into approximately one million X mg tablets and half a million 2X mg tablets, the remaining granulation was used to produce 4X mg tablets. This protocol was executed at two separate manufacturing sites (Sites A and B). According to the validation protocol the blends were considered uniform if they conformed to the USP Uniformity of Dosage Unit Criteria for finished tablets (i.e. individual units between 85% and 115% of label claim with an RSD less than or equal to 6%).

The final blending operation for these products is conducted in a large scale V-blender. Approximately half the contents of the blender is discharged into each of two transport hoppers after the mixing step is completed. These hoppers are then fed to a tablet press where the granulation is compressed into tablets of appropriate strength.

As part of validation, samples were withdrawn, with a thief, from six different locations within the V-blender just prior to discharge. The sampling sites, identified in FIGURE 1, included locations near the two trunnion regions of the blender. Five sets of samples (i.e. five thief stabs) were collected from each of the six locations in the order defined in FIGURE 1. In addition, fifteen single sets

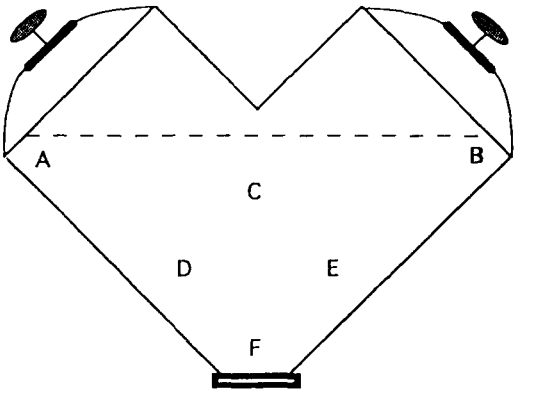

FIGURE 1: Sampling Locations (V-Blender).

of samples were collected from each hopper according to the sampling diagram presented in FIGURE 2.

A schematic diagram of the sampling thief, which was constructed in our machine shop, is presented in FIGURE 3. It is made of 316 stainless steel except for the handle, outer sleeve and sampling chamber inserts which are constructed of TEFLON™ to minimize binding. This thief was designed to remove three separate unit dose samples that contain X mg of DS and three separate samples that contain 2X mg of DS per stab. The six sampling ports are aligned vertically near the bottom of the thief; the three smaller ports are situated below the three larger ports. A sliding outer sleeve is used to cover the ports after the thief is removed from the bulk blend and is raised to reveal one port at a time as the samples are discharged into separate collection vials. The thieves used to sample the validation batches at the two manufacturing sites were different but identical in design.

One unit dose sample from each set at each location within the V-blender (a total of 30 samples) and one unit dose sample from each location within each hopper (15 samples per hopper) were submitted to our QC Laboratories for analysis. The entire sample was assayed for DS using an HPLC method, the same method used to analyze the final tablets. The remaining samples were saved as retains for further analysis as necessary.

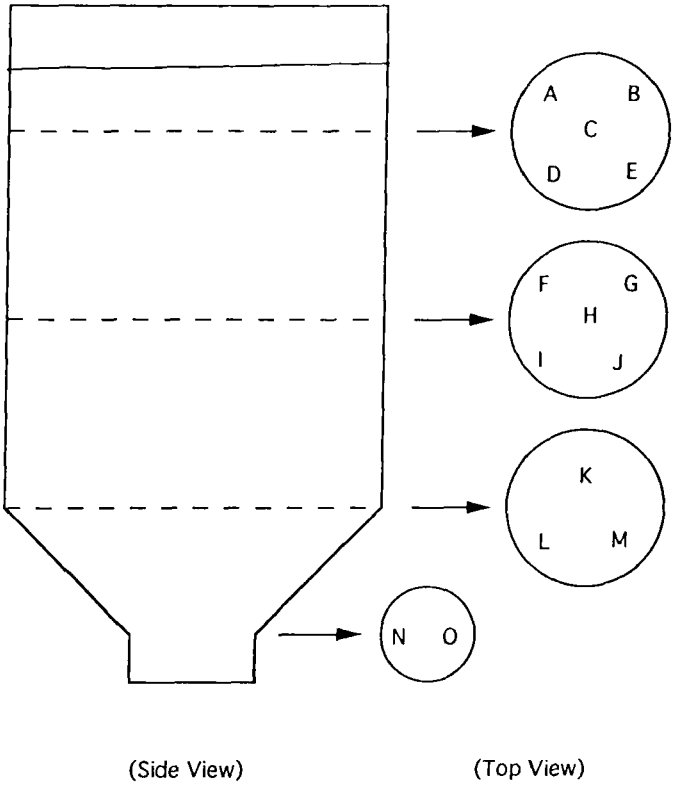

FIGURE 2: Sampling Locations (Transfer Hoppers).

RESULTS

Content uniformity data from each validation batch are presented in separate boxplots. Although a boxplot does not provide detailed information it can be used to quickly summarize the important characteristics of a particular data set. A boxplot consists of a box, range markers and perhaps individual data points. The box encloses 50% of the data; the median value of the variable displayed appears as a horizontal line within the box. The top and bottom boundaries of the box represent the upper (75^{th}) and lower quartiles (25^{th}) of the data population. The upper and lower quartiles are located halfway between the median and the highest and lowest values, respectively, of the data set. The range markers, which extend from the top and bottom of the box, indicate the extent of the main body of the

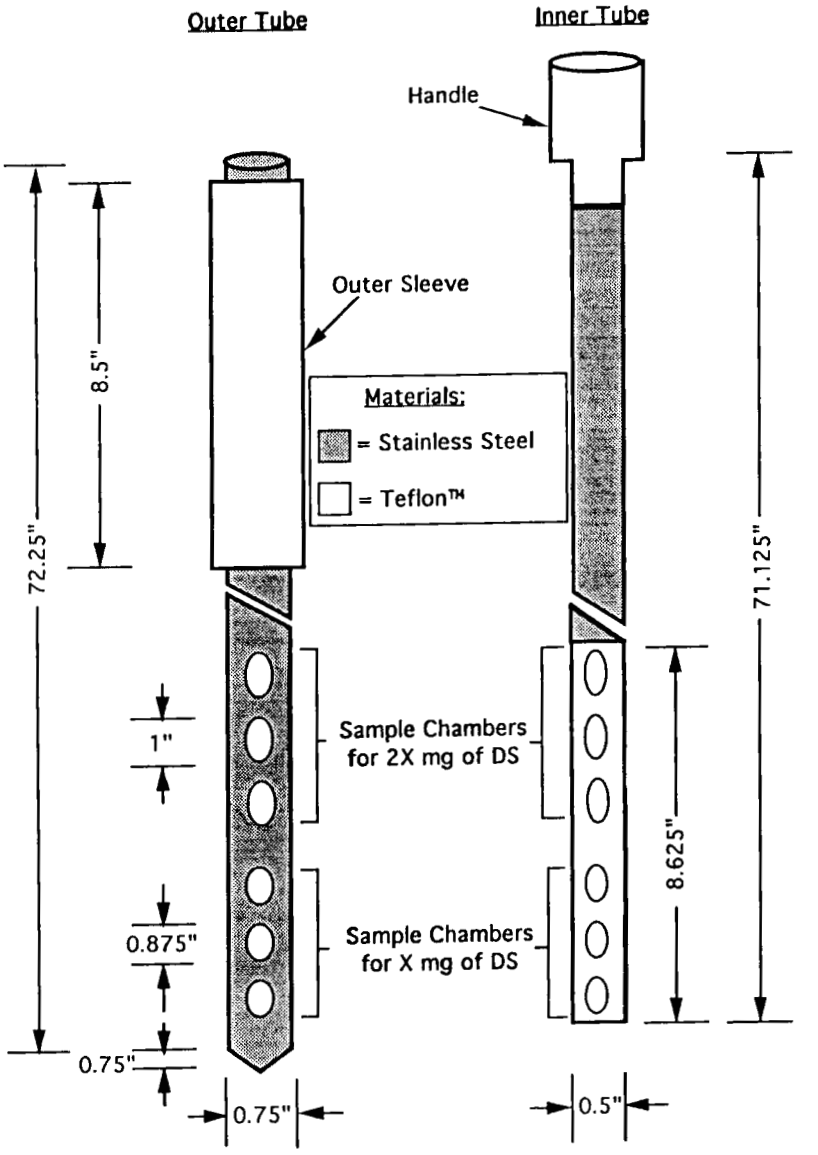

FIGURE 3: Sampling Thief (Not Drawn to Scale).

data. The length of the range markers are at most one and a half times the height of the box. Data that fall outside of the range markers are considered outliers and appear as individual data points. Since boxplots do not take into account sample size they cannot be used to determine if the differences between sets of data are statistically significant.

Boxplots that summarize the content uniformity data for the three successive validation batches manufactured at Site A are presented in FIGURES 4 - 6; similar results for the three batches manufactured at Site B are shown in FIGURES 7 - 9. Data are presented for unit dose samples (target weights that correspond to X mg of DS per sample) obtained from the V-blender and each of two transport hoppers. Content uniformity results are also shown for tablets that contain X mg, 2X mg and 4X mg of DS. The boxplots for each of the three batches manufactured at Site A (FIGURES 4 - 6) also show data collected from the V-blender at twice the target unit dose weight (corresponding to 2X mg of DS per sample). Sample sizes are indicated on the horizontal axis of each figure. The variation in the sample sizes between and within batches are due to differences in the number of retain samples that were submitted for analysis.

The data from the six validation batches exhibit some consistent trends which are summarized below:

1. The X mg, 2X mg and 4X mg tablets manufactured from each batch exhibit excellent content uniformity that readily conform to USP Uniformity of Dosage Unit Criteria for finished tablets. At least fifty X mg tablets (170 for Batch 1 Site A), collected during each compression run at both Sites were assayed. In addition, at least fifty 2X tablets (161 from Batch 1) were assayed from the three batches prepared at manufacturing Site A. The results, for the 4X mg tablets, are consistent with historical data from numerous commercial batches of this product.
2. With the exception of Batch 1 Site A, the RSD of all of the powder samples within a particular batch are comparable to the RSD of the corresponding tablets of the same strength.
3. With the exception of Hopper #2 from Batch 3 at Site A, the mean assay values for all of the powder samples are significantly lower than those of the corresponding tablets. In fact, only 36 (7.1%) out of 504 powder samples

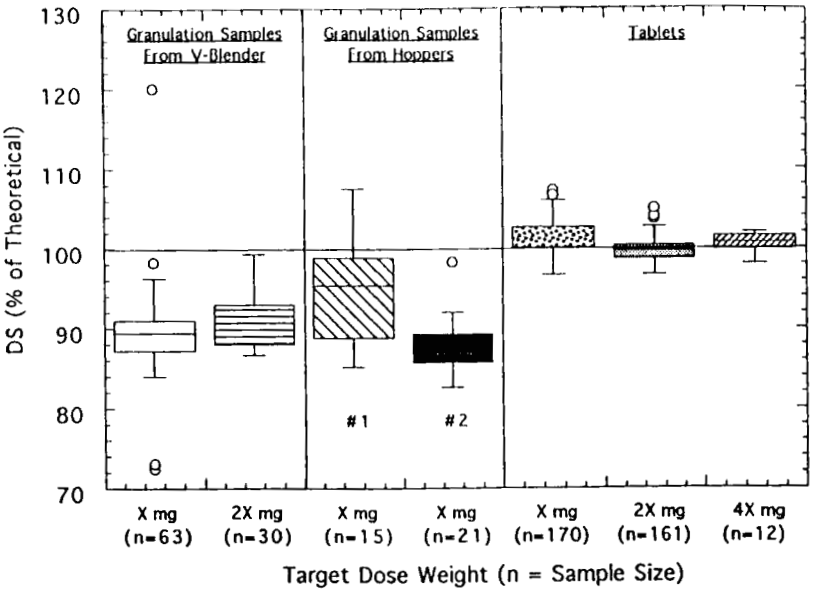

FIGURE 4: Content Uniformity of DS (Batch 1, Manufacturing Site A).

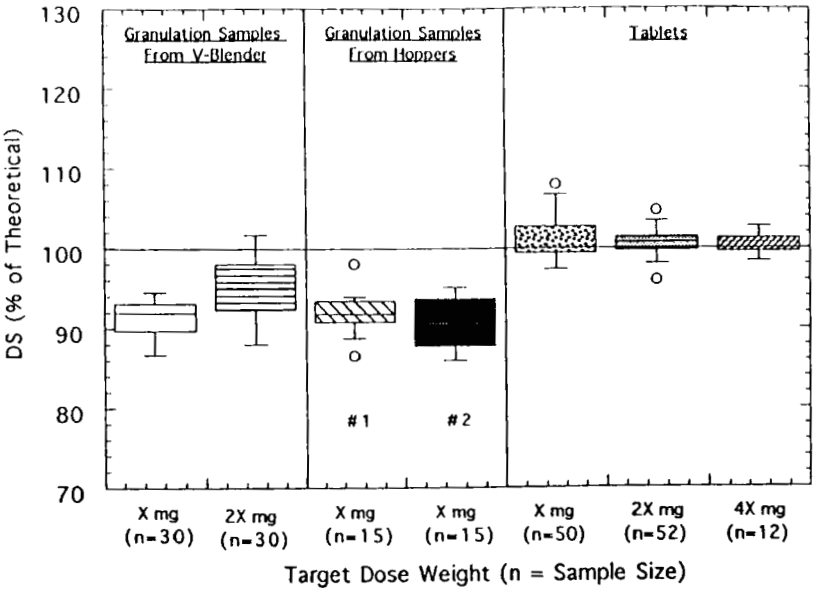

FIGURE 5: Content Uniformity of DS (Batch 2, Manufacturing Site A).

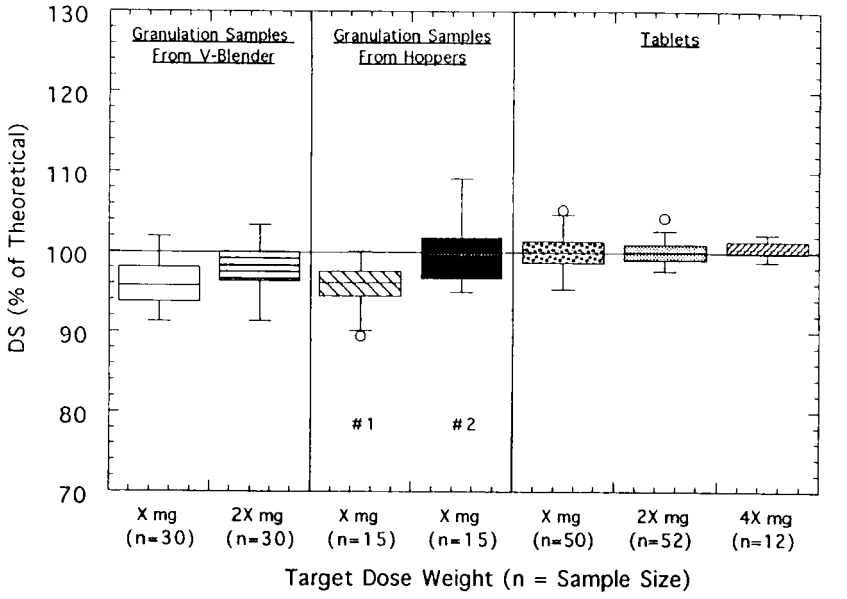

FIGURE 6: Content Uniformity of DS (Batch 3, Manufacturing Site A).

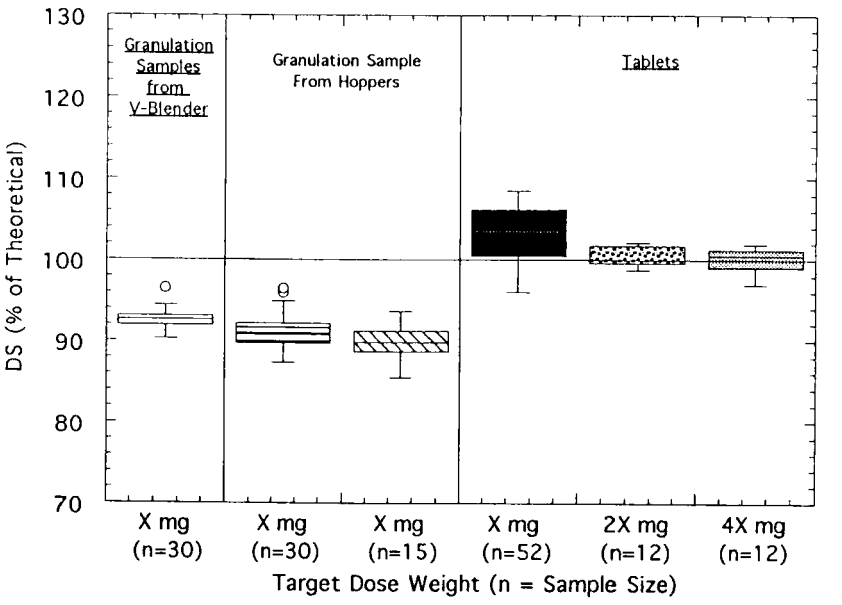

FIGURE 7: Content Uniformity of DS (Batch 1, Manufacturing Site B).

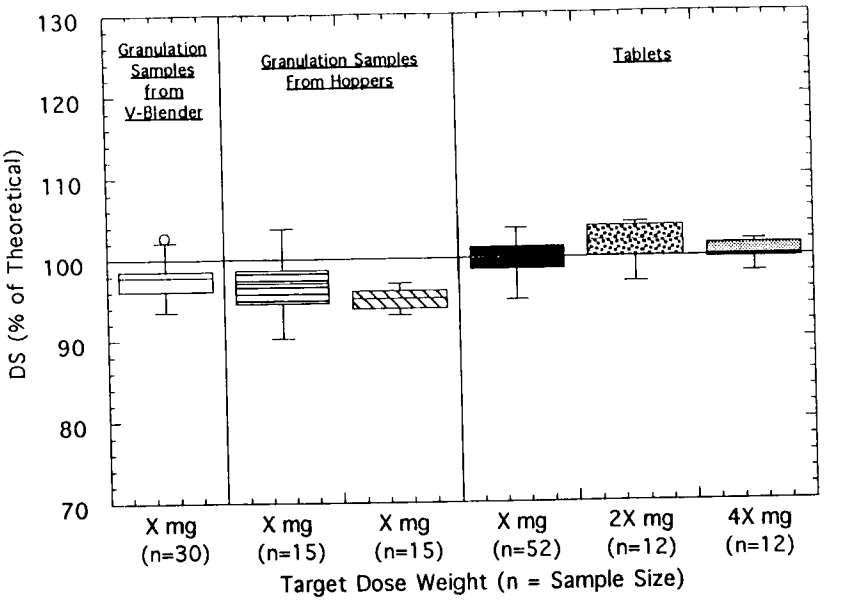

FIGURE 8: Content Uniformity of DS (Batch 2, Manufacturing Site B).

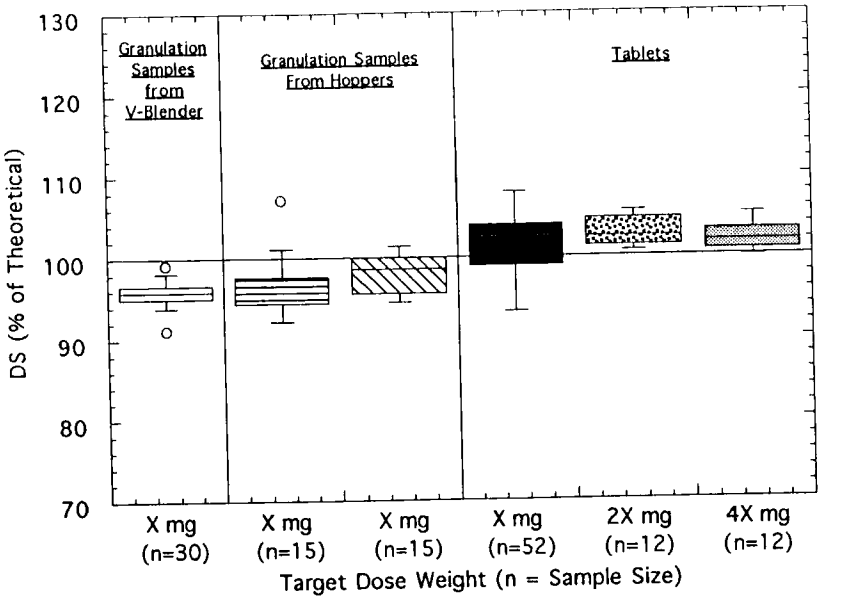

FIGURE 9: Content Uniformity of DS (Batch 3, Manufacturing Site B).

assayed from all six batches contain DS levels at or above 100% of label claim; of these, only 7 (1.4%) contain more than 103% of the theoretical quantity.

4. One of the batches (Batch 1 Site A) failed to meet the acceptance criteria specified in the validation protocol; two other batches (Batch 2 Site A and Batch 1 Site B) barely passed these criteria. Four out of the six batches would have failed the more restrictive content uniformity criteria (individual values between 90% and 110% with an RSD less than or equal to 6%) for powder blends that are suggested in recent FDA communications [2].

DISCUSSION

The key issue associated with this validation effort is not the variability of the data as characterized by the standard deviation but rather the consistently low assay results. It is these low values that prevented one of the batches from meeting the acceptance criteria specified in the validation protocol and would have prevented four of the batches from meeting the more demanding criteria that are outlined in recent FDA communications [2]. Although sampled from the same population, the mean DS content of the lubricated granulation is typically significantly lower than that of the corresponding tablets. In general, this could be caused by: (i) bias in the analytical method, (ii) enhanced blend uniformity as a result of unintentional mixing subsequent to discharge from the transfer hoppers and/or (iii) sampling bias. Since both the tablet and total powder samples were assayed by the same laboratory using an identical method it is unlikely that the differences in assayed DS content are due to bias in the analytical technique. The second possibility is more difficult to address other than to point out that a freely flowing powder with a non-uniform particle size distribution is more likely to segregate and become less rather than more uniform with additional handling.

The data suggest that the lower assay values of the powder specimens (relative to the tablets) may be due to sampling bias associated with the thief and is therefore not indicative of poor uniformity of the blend. After all, within a particular batch, the RSD's of the granulation samples are comparable to those of the corresponding strength tablets. Only one of the six validation batches exhibited unacceptably high RSD's for the targeted unit dose blend samples. This

was Batch 1 manufactured at Site A. As is evident from FIGURE 4, the large RSD of the specimens obtained from the V-blender for this batch is primarily due to three *widely* outlying data points. These outliers are significantly different from the finished product assays for this batch and all of the data from the other five validation batches.

These outliers include one super potent sample that was collected from location E (refer to FIGURE 1) and two subpotent samples that were obtained from position F. Interestingly, these two locations are located near the discharge valve of the blender, an area where poor mixing is not likely to occur. Additional testing of numerous retain samples from Batch 1 Site A was conducted in order to help determine the extent of this problem. The validation protocol specified that one targeted unit dose sample from each of five thief stabs at each of six locations within the V-blender be assayed for DS. This provided a considerable number of retain samples for further analysis since a total of 180 samples were withdrawn from the blender (each stab extracted three unit dose plus three twice-unit dose specimens). In the case of Batch 1 Site A, an additional 33 target unit dose samples were assayed for DS. Thirty of these retests included a second sample from each of the thief stabs. Three of these retests included the third sample from the three stabs that produced the original outlying data. Furthermore, one targeted twice-unit dose sample was assayed from each of the thirty stabs. All of the assay data from locations E and F for Batch 1 Site A are summarized in TABLE 1.

During validation five sets of samples were extracted from each of the nominal locations identified in FIGURE 1. Since the use of a thief disrupts the blend special care was taken not to insert the thief in the exact path made during a previous extraction. In addition, because the sampling chambers are arranged vertically on the thief (refer to FIGURE 3) all of the specimens from a particular stab were collected from a slightly different altitude within the blend. Thus, assays presented in TABLE 1 were conducted on samples that were not collected from identical locations in the V-blender. In light of these limitations it is impossible to draw definitive conclusions regarding the three outlying data points. However, based on the results presented in TABLE 1, the excellent uniformity of the tablets and the range of the data from all six validation lots, it is reasonable to assume that the three outlier assays may not be representative of undisturbed

TABLE 1

Assay Values (% DS of Theoretical) from Locations E and F in the V-Blender (Batch 1 Site A)

Location[1] - Stab[2]	X mg Target[3]			2X mg Target[3]
	Sample 1[4]	Sample 2[5]	Sample 3[5]	Sample 1[5]
E-1	87.2	87.0	-	90.0
E-2	88.3	90.1	-	91.0
E-3	90.5	91.5	-	91.7
E-4	90.8	90.5	-	91.7
E-5	120.1[6]	95.2	86.6	92.0
F-1	72.5	91.3	84.0	89.0
F-2	91.0	90.1	-	93.0
F-3	73.1	92.3	90.4	91.3
F-4	93.7	88.5	-	87.7
F-5	89.7	88.3	-	87.3

1 There are six sampling locations in the V-Blender (refer to FIGURE 1).
2 Five stabs are made per location.
3 Three samples at each targeted strength are extracted with each thief stab.
4 Normally assayed as per validation protocol.
5 Retain sample.
6 Outliers are underlined.
- Not assayed.

granulation residing in the V-blender. The RSD of the data would be reduced to an acceptable level if these three data points are excluded from the set.

One of the most troubling aspects of this validation effort was the large variability in the weights of the specimens collected from the V-blender and transfer hoppers. Ideally, the sampling thief should consistently remove volumetric aliquots of granulation that contain either X mg (unit dose) or 2X mg (twice-unit dose) of DS. The targeted sample weight is the weight of the

granulation that corresponds to these volumes. Unfortunately, the actual weights of all of the specimens collected during validation varied between 80% and 360% of these targets. The spreads of the normalized sample weights (actual weight divided by the target weight) are illustrated in FIGURES 10 - 12 for the targeted unit dose samples from Site A, the targeted unit dose samples from Site B and the targeted twice-unit dose samples from Site A, respectively. As is evident from these figures, the greatest variability is associated with the targeted unit dose samples from Batches 1 and 2 which were manufactured at Site A. The variability of the other data is much less pronounced.

The variability of the sample weights could be due to: (i) variation in the physical properties of the granulation, (ii) poor thief design and/or (iii) erratic sampling technique. Based on in-process particle size measurements and the excellent weight and content uniformity of the compressed tablets it is unlikely that the poor weight uniformity of the granulation specimens can be attributed to the first factor. We do believe, however, that the sampling chambers of the thief are larger than necessary. This accounts for normalized sample weights that are greater than unity but not for the variability of the data.

Circumstantial evidence suggests that the variability in the weights of the granulation specimens may be associated with the sampling procedure. Batches 1 and 2 manufactured at Site A were sampled by one pair of operators whereas Batch 3 was sampled by a second pair. All three of the batches manufactured at Site B were sampled by a third pair of operators. Furthermore, since instructions were not provided specifying the angle of insertion, orientation of the chamber, etc. it is probable that each pair of operators used different sampling techniques. It is clear from FIGURES 10 - 12 that the least variability occurred at manufacturing Site B and the most occurred during the first two batches at Site A. It is also apparent from these figures that the variability in sample weight for the twice-unit dose targets are significantly less than those of the unit dose targets (twice-unit dose data are not available from Site B). This would suggest that weight variation is reduced as the volume of the sampling chamber is increased.

The results presented in FIGURES 10 - 12 indicate a significant and important trend in the data. For a particular thief chamber volume, the percentage of DS in the sample decreases as the weight of the sample increases. If the blend

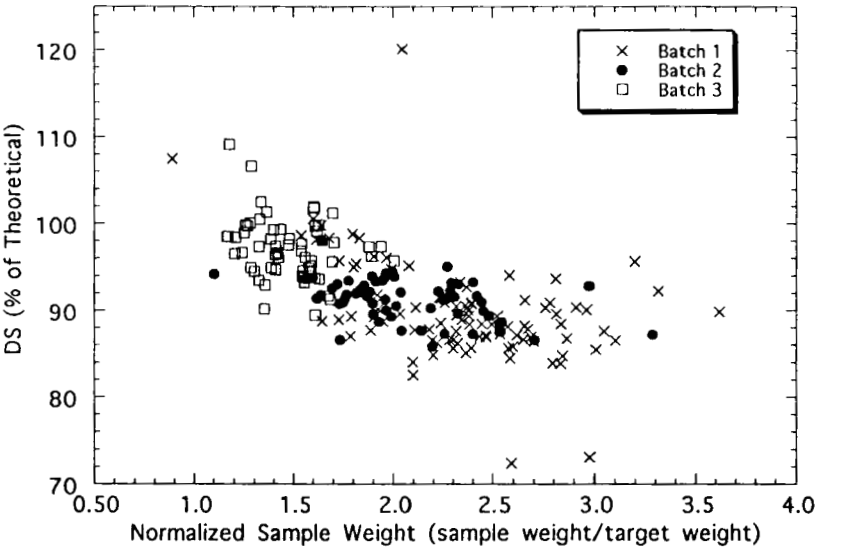

FIGURE 10: Assay vs. Sample Weight for Target Unit Dose Samples Manufactured at Site A (Combine V-Blender and Hopper Data).

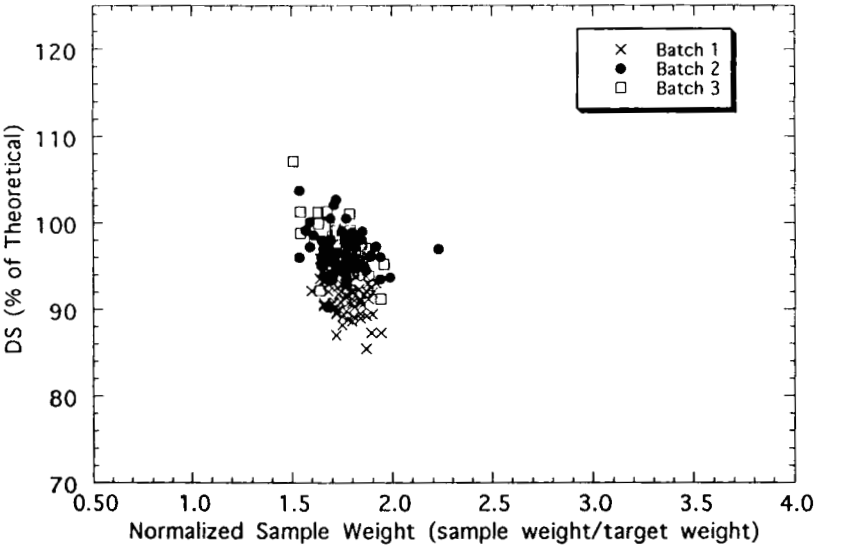

FIGURE 11: Assay vs. Sample Weight for Target Unit Dose Samples Manufactured at Site B (Combine V-Blender and Hopper Data).

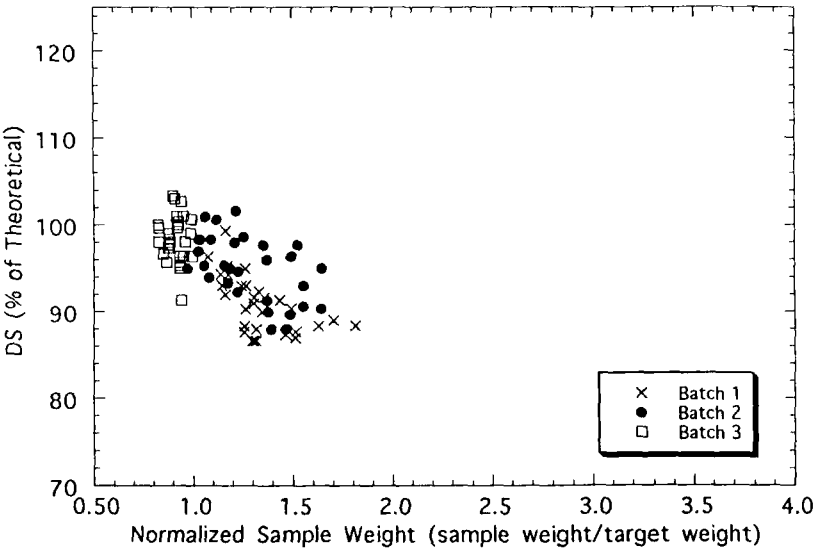

FIGURE 12: Assay vs. Sample Weight for Target Twice-Unit Dose Samples Manufactured at Site A (Combine V-Blender and Hopper Data).

was truly non-uniform at the X mg but not the 2X mg or 4X mg levels one would expect that larger sample weights would be associated with higher not lower assay values. This is because local non-uniformity would be masked in larger samples and the assay would approach that of the population which is obviously higher than the mean value of the targeted unit dose samples. The higher than expected normalized sample weights are likely due to "overfilling" of the thief chambers. Apparently, for this product, the thief chamber is more likely to be "overfilled" with excipient than with drug. This seems reasonable considering the lubricated granulation consists of relatively coarse drug containing granules plus a significant quantity of fine extra-granular excipients. Furthermore, the concentration of DS in the granulation is fairly low. One would expect that it is easier to "stuff" finer particles into a chamber of fixed volume than coarser particles.

Although technically a failure we feel that the results of this validation effort effectively demonstrate acceptable content uniformity of the final blend. This belief is based on: (i) excellent uniformity of large samples of three different

strengths of the finished dosage form, (ii) reasonable standard deviations of the powder blend samples and (iii) a long history of successfully and consistently manufacturing this product at the higher dosage strength.

As a result of these experiences we believe that alternative criteria need to be developed to determine the consistency of powder blends. Terms such as "assurance" and "consistently meet" are frequently encountered in FDA guidelines. Although these terms have not been fully defined they do suggest the need for statistically based acceptance criteria; this is a logical consequence for future guideline interpretation. It makes good business sense to evaluate validation data using sound statistical as well as scientific principles. Our firm has considered a number of statistically-based acceptance criteria. It is instructive to review these as it helps clarify the reasons for our eventual selection.

Single Factor Analysis of Variance

The Barr decision [1] requires the sampling of blender areas where weak mixing is possible. In V-blenders, areas adjacent to the trunnions have been so described [3]. One method of testing for poor mixing in the trunnion areas is to compare the average composition in these areas to other areas in the V-blender (refer to FIGURE 1). A meaningful sampling protocol should specify that four to six samples are taken from each area or zone. A single factor analysis of variance can then be used to determine if any one zone is statistically different from the others. Other well known tests can be used to establish which zone, or combination of zones, may be statistically different from the others. In principle, if the zones cannot be shown to be statistically different at a specified confidence level then acceptable uniformity can be assumed.

We found this criteria too stringent. Products which have longs records of excellent content uniformity, like the one described in this article, have routinely failed this test. In many cases areas of the blender other than the trunnion regions were shown to be significantly different from other, non-trunnion, zones. Batch to batch zone variance was also a factor. In one batch, variance within zones would be relatively small enabling acceptance of the hypothesis that the zones are significantly different. In the next batch, the zone variance might be large, and the hypothesis would be rejected. In both cases, the average concentrations from similar zones might be comparable.

Although we do not recommend analysis of variance for determining the acceptability of powder blends we continue to use the sampling scheme described above (refer to FIGURE 1). The determination that certain areas of a blender show a consistently high or low trend is valuable information indicating that further mixing studies may be warranted.

In the two methods which follow a different approach is taken. These methods involve acceptance criteria which relate to the ability of the population being tested to pass the USP content uniformity requirements; they both assume a normal distribution.

Bergum's Method [8]

Bergum [8] has published a method for constructing acceptance limits which relate to the ability of a population to pass the USP content uniformity requirements. While not in Bergum's original article, SAS® code exists for determining these limits.

The method first requires that a probability level be established for passing the content uniformity test. For example, if one chooses a probability level of 95%, a region is defined in the standard deviation - mean composition space (of the **population**) at which the probability is 95% or higher that subsequent **samples** from the population will pass the test. Any combination of **population** mean and standard deviation falling within this space is guaranteed at least a 95% probability of passing the multistage test. Since the mean and standard deviation of the population are estimated from samples one must next determine a joint confidence region for the **sample** standard deviation and mean. This assures, at a specified confidence level, e.g. 90%, that the **sample** was taken from a **population** within the region determined above. Bergum used a Monte Carlo technique to validate his method. This simulation indicated a population with a potency of 100% of label claim and a 6% standard deviation has a 95.7% probability of passing the USP content uniformity requirements for tablets.

The acceptance region for the sample will be dependent upon the probability level and the confidence limits chosen plus the sample size. TABLES 2 and 3 provide examples of these regions for tablets and capsules, respectively. For a given mean and sample size, the %RSD of the sample must be less than or equal

TABLE 2

Critical Values of the %RSD for Passing USP Content Uniformity Requirements for Tablets[1]
(Based upon a Method Proposed by Bergum [8])

Number of Samples, n	Mean Assay (% of Label)				
	90	95	100	105	110
10	1.18	2.23	3.17	2.02	0.96
20	1.50	2.85	4.01	2.58	1.23
30	1.66	3.15	4.39	2.85	1.35
40	1.76	3.33	4.61	3.01	1.44
50	1.83	3.45	4.75	3.12	1.49
60	1.88	3.55	4.86	3.21	1.54
70	1.92	3.63	4.94	3.29	1.56
100	2.00	3.78	5.11	3.42	1.64

[1] A %RSD of the sample less than or equal to the critical %RSD provides a 90% assurance that at least 95% of all samples tested for content uniformity will pass the USP test for tablets.

to the indicated value for the test to be successful. Note in TABLE 2, for a sample size of 10 tablets and a mean of 100% of label claim, an RSD of 3.17% or less is required for acceptance. If the sample size is increased to 30, an RSD of not more than 4.39% is required. Obviously, these requirements are considerably more restrictive than the USP test.

The Bergum method is an excellent method for demonstrating the ability of a batch of tablets or capsules to meet content uniformity requirements. We have adopted it for use in our tablet and capsule content uniformity validation efforts. Our experience in the blender validation effort presented in this paper, and in other similar blender validation efforts, is that failures will occasionally occur when this method is applied to unit dose data. Furthermore, these failures have

TABLE 3

Critical Values of the %RSD for Passing USP Content Uniformity Requirements for Capsules[1]
(Based upon a Method Proposed by Bergum [8])

Number of Samples, n	Mean Assay (% of Label)				
	90	95	100	105	110
10	1.48	2.79	3.67	2.52	1.21
20	1.92	3.63	4.51	3.29	1.57
30	2.13	4.03	4.87	3.65	1.75
40	2.27	4.28	5.08	3.88	1.85
50	2.37	4.46	5.22	4.04	1.94
60	2.44	4.59	5.32	4.16	2.00
70	2.50	4.69	5.41	4.26	2.05
100	2.61	4.91	5.57	4.46	2.14

[1] A %RSD of the sample less than or equal to the critical %RSD provides a 90% assurance that at least 95% of all samples tested for content uniformity will pass the USP test for capsules.

occurred in products with long histories of excellent content uniformity performance and in which extensive testing of the product in the validation batches indicated excellent uniformity. As suggested in this paper, we believe these failures are attributable to sampling bias rather than poor uniformity. We have therefore discontinued the use of this criteria for blender validation.

Standard Deviation Prediction Interval

The primary measure of blender uniformity should be the standard deviation of the active (or actives) within the blend rather than the %RSD. Our experience

has shown that thief sampling of stagnant beds sometimes results in mean potency measurements significantly different from those of the final dosage form and from the theoretical potency as obtained by validated weigh-out procedures. This tends to bias the %RSD which is the standard deviation normalized to the mean.

The standard deviation has been used extensively as a measure of blend uniformity in the powder technology literature [9]. Mixing theory has shown that a random blend of dissimilar particles will have a non-zero minimum standard deviation. We recommend the standard deviation prediction interval as an appropriate acceptance criteria when used in conjunction with other tests.

The standard deviation prediction interval is discussed by Hahn and Meeker [10] and may be easily calculated with the use of a table of F values. In general, the method allows one to predict from a sample of size n, with specified confidence limits, an upper bound for the standard deviation of a future sample of size m from the same population:

$$S_{cr} = S_m / [F_{1-\alpha,m-1,n-1}]^{1/2} \quad (1)$$

in which:

S_{cr} = critical value of the standard deviation from a sample of size n,
n = sample size during validation,
S_m = upper bound of standard deviation of a subsequent sample of size m,
m = the size of a subsequent sample (= 10 for USP Stage 1 testing),
$1-\alpha$ = confidence interval (e.g = 0.9) and
F = the F distribution.

When discussing the Bergum method we noted that a population RSD of 6% of the target concentration was shown to provide a high degree of probability (95.7% when the potency is 100%) that the sample would pass the USP test for content uniformity. A 6% RSD or less is also required to pass the first stage of the USP test. Thus we recommend the upper bound of the standard deviation be:

$$S_m = S_{10} = 0.06 \times [\text{target concentration}].$$

The standard deviation prediction interval allows us to calculate the maximum (i.e. critical) acceptable standard deviation for unit dose samples of a final blend taken during validation. For the acceptance criteria, a maximum standard deviation is specified. A 90% confidence interval is recommended in making the calculation. This critical standard deviation guarantees, with an

TABLE 4

Example Calculation - Standard Deviation Prediction Interval for the Three Validation Batches Manufactured at Site A (Data from V-Blender)

$1-\alpha = 0.9$ $m = 10$ $S_{10} = 6.0$				
Sample Size, n	Standard Deviation, (% of Label)	$F_{1-\alpha,m-1,n-1}$	Critical Std Dev[1] (% of Label)	Pass /Fail Criteria
63	5.752 (Batch 1)	1.735	4.555	Fail
30	2.200 (Batch 2)	1.857	4.403	Pass
30	2.797 (Batch 3)	1.857	4.403	Pass

[1] Calculated from Equation (1)

assurance of 90%, that the upper prediction bound for a future sample of size 10 will not be greater than 6% of the target concentration. For example, the standard deviation of the targeted unit dose blender samples from the three validation batches conducted at site A are presented in TABLE 4.

Note that in the above example the V-blender samples from the first batch manufactured at Site A failed to pass the test. This same batch also failed this criteria when based on the hopper data. All three batches manufactured at site B passed this acceptance criteria (based on both blender and hopper). One should note that this criteria is considerably more stringent than the USP test for content uniformity. Here, for a sample size of 30, a standard deviation of 4.403% or less is required for acceptance. This contrasts with the USP criteria of 6.0 for a sample size of 10 at the first stage.

Potency is established by appropriate and validated weigh-out procedures, stability of the active(s) during processing, and safeguards against either dilution or concentration of the active(s) during manufacturing. Obviously, one needs to establish potency in the final dosage form in addition to that of the final blend. We therefore advocate extensive testing of the final tablet or capsule product during blender validation. This not only establishes that acceptable potency levels are met, but also demonstrates the uniformity measured in the blender was not compromised by particle segregation in post blending operations. The Bergum method, discussed above, is used to establish a high degree of statistical confidence that a future sample of the **finished product** will meet USP requirements for content uniformity. Thus, we advocate the use of the standard deviation prediction interval for the final blend and the Bergum **[8]** method for the finished product.

CONCLUSIONS

Prior to executing the validation protocol we felt, perhaps naively, that demonstrating content uniformity of a powder blend was a relatively straight forward exercise. We now feel otherwise. Our validation effort for this product failed even though a large number of samples of the finished dosage form exhibited excellent content uniformity. It failed primarily because the drug content of some of the targeted unit dose granulation samples were below the required 85% lower bound. We believe that this was due to sampling bias and, as a result, the blend specimens that were assayed were not representative of the population. We speculate that sampling bias, which was manifested as highly variable specimen weights, was due to "overfilling" of the thief chamber which in turn, at least for this product, led to lower concentrations of drug in the samples than in the population. Consequently, the RSD's of the samples were biased on the high side since the sample means were biased on the low side.

These data suggest, as other have concluded **[7]**, that content uniformity of a final blend is best demonstrated, for some formulations, by sampling and assaying entire blend specimens that are larger than unit dose. Unfortunately, the Barr **[1]** decision essentially eliminates this as an option for manufacturer's of pharmaceutical products. In light of this restriction we feel that emphasis should

placed on minimizing sampling bias that will inevitably occur when small volume samples are extracted with a thief from large volume populations. This goal can be approached with proper thief design and careful and consistent sampling technique.

Another critical aspect of process validation concerns the criteria used to assess the uniformity of powder blends. These criteria should be based on sound statistical principles. We recommend a two pronged approach in which the Standard Deviation Prediction Interval is applied to unit dose samples from the final blend and the Bergum [8] criteria is applied to the finished product. Both of these criteria are more stringent than the USP test for content uniformity.

ACKNOWLEDGMENTS

The authors wish to thank: John Shelton, Dave Whiteman and Jerry Lewis for their valuable input concerning the statistical criteria used in this study; and Phil Schneider and Aaron Schoeneman for supervising the execution of the validation protocols.

REFERENCES

[1] 812 F. Supp. 458 (District Court of New Jersey, 1993).

[2] FDA, "Guide to Inspections of Oral Solid Dosage Forms Pre/Post Approval Issues for Development and Validation," January, 1994.

[3] "Tips for Improved Mixture Sampling," Powder and Bulk Engineering, 7(1), 41 - 45 (Jan 1993).

[4] T. Allen, "Particle Size Measurement, Fourth Edition," Chapman and Hall, London, 1990.

[5] R.L. Lantz Jr. and J.B. Schwartz in "Pharmaceutical Dosage Forms - Tablets, Volume 2, Second Edition," H.A. Lieberman, L. Lachman and J.B. Schwartz, eds., Marcel Dekker, New York, 1989, pp. 27 - 32.

[6] R.L. Lantz Jr., in "Pharmaceutical Dosage Forms - Tablets, Volume 2, Second Edition," H.A. Lieberman, L. Lachman and J.B. Schwartz, eds, Marcel Dekker, New York, 1989, pp. 158 - 162.

[7] J.T. Carstensen and C.T. Rhodes, Drug Development and Industrial Pharmacy, 19(20), 2699 - 2708 (1993).

[8] J.S. Bergum, Drug Development and Industrial Pharmacy, 16 (44), 2153 - 2166 (1990).

[9] F. Pitard, "Pierre Gy's Sampling Theory and Sampling Practice, 2nd Ed.," CRC Press, Boca Raton, FL, 1993.

[10] G.J. Hahn and W.Q. Meeker , "Statistical Intervals, a Guide for Practitioners," John Wiley & Sons, New York, 1991.