



- 34           – at each design point, this procedure is repeated 1000 times to get a distribution of  
35           90<sup>th</sup> (or 95<sup>th</sup>) percentiles.
- 36           • Some possible distributions of the 90<sup>th</sup> percentile are illustrated on the last page of  
37           Exhibit A. At point A, both Test and Reference variances are low, and at point B, the  
38           Test variance is high while the Reference variance is low. If the distributions of 90<sup>th</sup>  
39           percentiles at A and B do not overlap, as shown in the illustration, the method has  
40           good discrimination between these two situations (A and B), and will make a good  
41           decision-making tool if indeed one wants to distinguish between these two  
42           situations. Note that the distribution could be broad (as at design point C), yet be  
43           shifted sufficiently far from the other scenarios so that the method would still have  
44           good discrimination. Whether the chi-square actually behaves the way presented in  
45           the illustration, remains to be seen.
  - 46           • At the next step, if the Working Group decided that situations A and B should be  
47           judged equivalent but situation C inequivalent, then the critical value would have to  
48           be set around 11.5 (using the illustration's values).
  - 49           • In addition to the high percentile(s), other percentiles (e.g., 50<sup>th</sup>) could be calculated  
50           to assess properties of the chi-square ratio distribution; for instance, it may be found  
51           that 50<sup>th</sup> percentile is more stable than 90<sup>th</sup> or 95<sup>th</sup>. (The additional percentiles will  
52           not add to the total calculation time.) The lower percentiles (e.g., 25<sup>th</sup>) are of less  
53           interest because they represent situations of low chi-square ratios, i.e., the area  
54           where profiles are equivalent and differences between profiles are small. The  
55           regulatory decisions will be made around high chi-square ratios.
  - 56           • For simplicity, the within-lot variability was assumed equal to between-lot in this  
57           protocol.
  - 58           • Initially, only a limited number of scenarios will be studied. If the method behaves  
59           well there, other situations might be studied. The protocol might be revised to list  
60           the design combinations which might be studied beyond the initial ones.

61           The participants agreed that the protocol should be slightly revised to include: (i) the  
62           definition of "scenario"; (ii) a description of how simulations are conducted; (iii) a section on  
63           how the work will be reported; and (iv) directions for future work. Eventually, SAS  
64           programming code will be included in the protocol. The Working Group agreed that further  
65           fine-tuning of the protocol should be done after FDA provides detailed comments on the  
66           current draft with the slight revisions listed above.

### 67           Draft Paper

68           Dr. Lyapustina explained that the PQRI publication policy recommends that work of  
69           PQRI Working Groups be submitted for publication to the journals of the member organizations  
70           as a first choice. Since the International Journal of Pharmaceutics is affiliated with AAPS, it is  
71           considered an appropriate publication source for the Working Group's paper.

72           She reviewed the progress of the paper as follows:

- 73           • Introduction and background has been drafted by Dr. Adams and his FDA  
74           colleagues.

- 75 • The review of the chi-square method is to be drafted by Dr. Tsong, possibly in July.
- 76 • The overview of the early work of this Working Group, and in particular of the work  
77 done by Dr. Lee and Dr. Pan, will likely be drafted in August.
- 78 • The future direction of the work will be described based on the protocol being  
79 developed by the Working Group currently, under the leadership of Drs. Morgan,  
80 Pan and Mr. Christopher.

81 The participants agreed that the details of the current work should be published as a  
82 separate article. The interim progress report will outline this work only in general terms. Both  
83 articles could be referenced in the final recommendation to be submitted to FDA.

#### 84 **IV. AGREED**

- 85 • The draft protocol will be slightly revised (see Exhibit A) and re-circulated for  
86 further comment by the Working Group, especially by FDA representatives.

#### 87 **V. NEXT MEETING/TELECONFERENCE**

88 The next teleconference will be held on 21 July if sufficient feedback on the revised  
89 protocol is received by then. Otherwise, the next teleconference will be held in August.

90 Finalized on 9 September 2004

#### 91 **VI. ADDENDUM**

92 After the teleconference, Dr. Pan and Mr. Christopher clarified the following by email:

93 The component of a single stage in the denominator of the chi-square formula is  $(R1-R2)^2 / [(R1+R2)/2]$ .  
94 It can be rewritten as  $2 [(R1-R2)/(R1+R2)] (R1-R2)$ .

95 Using limit theory, when R1 and R2 both approach zero,  $(R1-R2)$  (the last part in the rewritten formula) is  
96 approaching zero. The value of  $[(R1-R2)/(R1+R2)]$  is limited within the range  $[-1, 1]$  conditional on  $R1 > 0$   
97 and  $R2 > 0$ . So the whole thing approaches zero when both R1 and R2 approach zero from the positive  
98 side. This could be used as the justification that zero is actually a reasonable value to be used for a given  
99 stage if both R1 and R2 have zero deposit on it.

100 From the original question, if both R1 and R2 are zero, then  $R1+R2$  is zero, thus  $(R1-R2)^2 / [(R1+R2)/2]$   
101 is not defined since both the denominator  $(R1+R2)/2$  and numerator  $(R1-R2)^2$  are zeros. But from the  
102 above, the whole thing is approaching zero as long as both R1 and R2 are approaching to zeros from  
103 positive side which is the case in our application.

104 If R1, R2, and T for a given stage are all zeros, then we have  $[(R1+R2)/2 - T]^2 / [((R1+R2)/2+T)/2]$  as  
105 undefined. Let Ravg replace  $(R1+R2)/2$ , then we can rewrite  $[(R1+R2)/2 - T]^2 / [((R1+R2)/2+T)/2]$  as  
106  $(Ravg - T)^2 / [(Ravg+T)/2]$ . So it is reduced to the form  $2 (x-y)^2 / (x+y)$ . By the same argument, we can  
107 conclude that its value is zero if **all** R1, R2, and T are approaching zero from positive side. The situation  
108 when R1, R2 and T are all zero could only occur on one or two stages but not on all stages. (Otherwise  
109 the run would be invalid. Only valid CI runs should be used for profile comparisons). After summation,  
110 the whole denominator, namely the summation of  $(R1-R2)^2 / [(R1+R2)/2]$  over all of the stages, cannot  
111 be zero if two reference CI runs are not identical.

112 So considering the original formula, the only situation we need to handle is that both reference CI runs  
113 are identical. If this situation can be safely avoided, then we should be fine in terms of the definition of the  
114 chi-square statistic.

115 **VII. EXHIBIT A**

116 **Initial Protocol to Investigate Properties of Chi-Squared Method**

117 *(Draft as clarified per the 7 July teleconference discussion, version 4.1, dated 13 July 2004)*

118 **I. Objective**

- 119 • Evaluate the Chi-Squared approach, based on bracketing a range of changes in the mean  
120 and variability, for particle size distribution profiles for a specific product type (e.g., CFC  
121 suspensions, HFA suspensions, dry powder inhalers).

122 **II. Methodology**

123 **A. Define the scenarios to include.**

- 124 • Initial investigation will be based on combinations of changes in the mean profile of the test  
125 population and changes in the levels of test and reference variability.

126 **1. Changes in the Test Population mean profile**

127 The levels of this factor to be studied will be:

- 128 1. No change  
129 2. 10% increase for 1 site with high recovery  
130 3. 10% increase for 1 site with medium recovery  
131 4. 10% increase for 1 site with low recovery  
132 5. 100% increase for 1 site with high recovery  
133 6. 100% increase for 1 site with medium recovery  
134 7. 100% increase for 1 site with low recovery

135 Definitions of high, medium, and low recovery sites will be dependent on product type  
136 and will be guided by analysis of industry data + additional sources.

137 Recoveries on the other stages and sites will decrease proportional to the percent  
138 deposited at that site.

139 **2. Changes in the levels of test and reference population variability**

140 Define a 'Mid' level of variation to be the estimated RSD associated with a  
141 particular site from the observed data profiles. The 'Low' level represents **half** the  
142 estimated RSD from each site, while the 'High' level represents a **tripling** of the  
143 estimated RSD. The estimated RSD associated with a particular site reflects the total  
144 variability from the observed data.

145 The total number of combinations studied would be:

- 146 1. Reference variation Low, Test variation Low  
147 2. Reference variation High, Test variation Low  
148 3. Reference variation Low, Test variation High  
149 4. Reference variation High, Test variation High  
150 5. Reference variation Mid, Test variation Mid

- 151 • Based on this design in which the 5 levels of test and reference population variability are  
152 paired with the 7 levels of changes in the mean profiles, a total of **35** scenarios will be  
153 considered. Three more additional combinations will be added to investigate midpoint  
154 changes: 50% increase in the test population mean with mid levels of test and reference  
155 variability across low, medium, and high recovery sites. This brings the total scenarios to **38**.  
156 A visual representation of the design space and the chosen combinations is given on page 3.

- 157
- 158
- 159
- Summary statistics (means and standard deviations) and graphical profiles of the baseline data for CFC suspensions are presented in Figure 1 (Figure 1 included as a separate document).

160 **B. *Conduct simulations for each scenario.***

- 161
- 162
- 163
- 164
- 165
- 166
- Simulations will incorporate inter-stage correlation observed from absolute recovery data.
  - The variation in the simulated data will bracket the variation observed from the baseline data.
  - For each scenario or design point, the chi-square algorithm will be applied 1000 times. This means that 30 Test and 30 Reference profiles will be generated 1000 times, and each time the profiles are generated, the data will be simulated as if the units were tested from one batch; i.e., the total variation is the within batch variability.
  - For each simulation, the 30 Test and 30 Reference profiles will yield a population of unique triplets of profiles (Test, Ref1, Ref2). From this population, 500 triplets will be sampled and the mean ratio of chi-squares obtained. The selection of 500 triplets and calculation of the mean ratio will be repeated 300 times, yielding, for each simulation, a distribution of 300 means. From this distribution of 300 means, 4 summary statistics will be reported: the 50<sup>th</sup> percentile, the 90<sup>th</sup> percentile, the 95<sup>th</sup> percentile, and the mean. This entire process will be repeated 1000 times, resulting in a distribution of 1000 90<sup>th</sup> percentiles, 1000 95<sup>th</sup> percentiles, etc. The attached Appendix contains a more detailed summary of the implementation of the algorithm.

176 **C. *Report results.***

- 177
- 178
- For each of the four summary statistics listed above, the entire distribution across the 1000 simulations will be reported

179 **III. Assessment**

180 **A. *Stability Assessment of Chi-Square Algorithm***

- 181
- 182
- 183
- The assessment will primarily focus on understanding how changes across the design space influence the location and shape of the distributions resulting from 1000 applications of the chi-square algorithm.

184 This increased understanding will:

185 (1) highlight additional areas within this design space (or more

186 complicated design spaces) that need further investigation

187 (2) eliminate certain regions of the design space for further investigation.

- 188
- 189
- The information gathered from this initial protocol will help determine what designs should be considered next when investigating changes across multiple impactor deposition sites.

190 **B. *Procedure for Determining Critical Value***

- 191
- 192
- 193
- Secondly, the assessment will explore options for determining critical values. Selection of a critical value should incorporate understanding both the location and shape of the distributions resulting from the simulations.

194 **C. *Define Further Work (Left blank at this point pending comments by FDA***

195 ***representatives)***

196

197

198 **IV. Implementation of Chi-Square Algorithm**

199 For each simulation at a particular design point, the following steps will be followed in implementing the  
200 chi-square algorithm:

- 201 1. From the 30 reference product profiles and 30 test product profiles, randomly select with  
202 replacement, 500 “triplets,” each triplet consisting of two reference product profiles (Ref#1 and Ref#2  
203 where Ref#1 and Ref#2 are different profiles) and one test product profile (Test)
- 204 2. For each triplet calculate the following for each stage of the profile (i.e., throat, filter, plate1, plate 2,  
205 etc.):
- 206 ➤ **Expected Ref** = (Ref#1 + Ref#2) / 2
  - 207 ➤ **Ref, Ref Difference** = Ref#1 – Ref#2
  - 208 ➤ **Test, Ref Difference** = Test – Ref
  - 209 ➤ **Expected Test, Ref** = (Test + Ref) / 2
  - 210 ➤ **Test, Ref Deviation** = (Test, Ref Difference)<sup>2</sup> / Expected Test, Ref
  - 211 ➤ **Ref, Ref Deviation** = (Ref, Ref Difference)<sup>2</sup> / Expected Ref
- 212 3. For each triplet calculate:
- 213 ➤ **Chi-sq(Test:Ref)** = Sum of Test, Ref Deviations for all stages
  - 214 ➤ **Chi-sq(Ref:Ref)** = Sum of Ref, Ref Deviations for all stages
  - 215 ➤ **Chi-square Ratio** = Chi-sq(Test:Ref) / Chi-sq(Ref:Ref)
  - 216 ➤ This can be depicted in equation form as the following :

217

$$Chi - square Ratio = \frac{\sum_{i=1}^n \left( \frac{(T_i - \bar{R}_i)^2}{\frac{T_i + \bar{R}_i}{2}} \right)}{\sum_{i=1}^n \left( \frac{(R_{i,1} - R_{i,2})^2}{\frac{R_{i,1} + R_{i,2}}{2}} \right)}$$

218 where  $\bar{R}_i = \frac{R_{i,1} + R_{i,2}}{2}$

219 and  $n$  = the number of deposition sites of the profile

- 220 ➤ **NOTE 1:** If  $R_{i,1}$  and  $R_{i,2}$  are both 0 at a particular site, then the component associated with that  
221 particular site from the sum in the denominator will be assigned a 0.
- 222 ➤ **NOTE 2:** The algorithm will be applied to data calculated in terms of percent impacted.

223 4. Then calculate:

- 224 ➤ **Mean Chi-square Ratio** = Mean of the 500 Chi-square Ratios

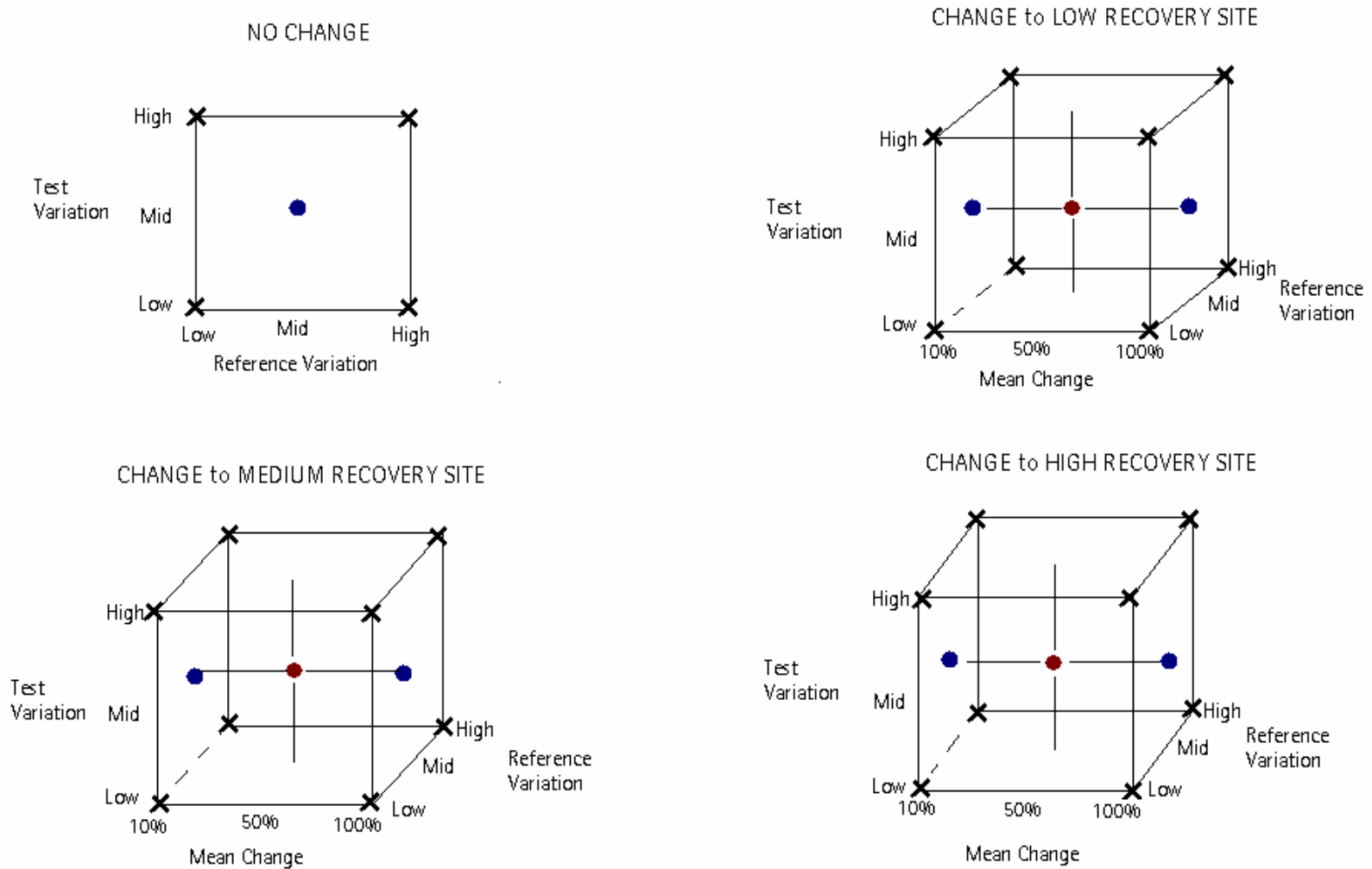
225 5. Repeat steps 1 through 4 300 times to obtain a distribution of 300 Mean Ratios

226 6. From this distribution of 300 Mean Ratios, report 4 summary statistics : the 50<sup>th</sup> percentile, the 90<sup>th</sup>  
227 percentile, the 95<sup>th</sup> percentile, and the mean

228 7. Repeat steps 1 through 6 1000 times.

229

### Visual Representation of Design Combinations for Initial Protocol



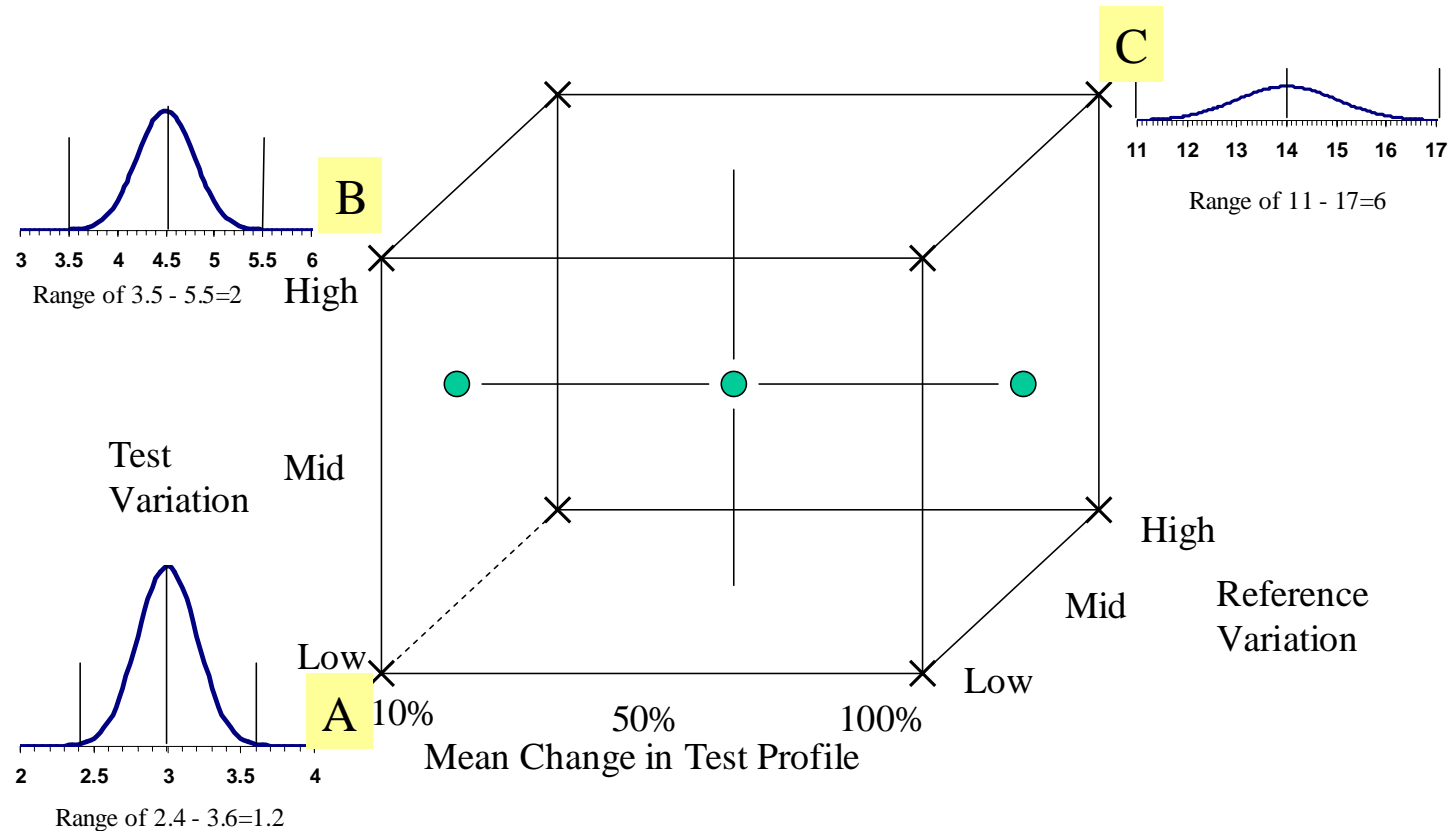
230

231

232

233

### Possible\* Distribution of 1000 90th percentiles\*\*



\* The drawn distributions are “made up,” have no relation to the actual chi-square method, and are shown only for illustration of the concept.  
 \*\* Each 90th percentile would be calculated from a distribution of 300 sample means.

234