



Biopharmaceutical Emerging Best Practices Association

PO Box 7087
Citrus Heights, Ca 95621
916.729.0109 v
916.729.2602 f
<http://www.bioquality.biz/bebpa/index.htm>

To: Tina S. Morris, Ph.D
Senior Scientist
Scientific Liaison, Council of Experts Committee for Vaccines, Virology, and Immunology Complex Actives Division Department of Standards Development U.S. Pharmacopeia
12601 Twinbrook Parkway
Rockville MD 20852-1790
Tel.: 301-816-8397
Fax: 301-816-8373
e-mail: tsm@usp.org

USP Biostatistics Expert Committee Ad Hoc Panel for the Revision to USP General Chapter Analysis of Biological Assays <111>

The Biopharmaceutical Emerging Best Practices Association (BEBPA) is writing to submit comments on the Revision to the USP General Chapter Analysis of Biological Assays <111> (Fall 2007 version). These comments were collected at the Inaugural BEBPA Bioassay Conference held in Berlin September 10-12, 2008 during a workshop on parallelism and two discussion sessions dedicated to the GC <111> Revision.

BEBPA (www.bioquality.biz/bebpa/) is a not-for-profit association which aims to provide an international forum for the presentation and discussion of scientific issues and problems encountered in the biopharmaceutical community. Approximately 80 individuals, representing large and small biopharmaceutical manufacturing companies, clinical research organizations, consultants and government agencies, with experience and interest in a wide range of biological pharmaceutical products, supported by a variety of biological assays, attended the conference. All conference attendees were invited to participate in the discussions. To facilitate free discussion it was announced that comments would not be recorded as being attributed to any individual. This was done to permit the expression of personal opinions which might differ from the official positions of the organizations to which the participants were affiliated. Notes were taken by several volunteers. The discussion was animated, and this letter aims to convey the major concerns expressed. It should be noted the intent was not to achieve a consensus amongst all attendees and the points below were not necessarily unanimously agreed upon. Furthermore, it should be noted that some participants indicated that they had, or intended to, write independently to the USP about the issues on which they felt strongly.

There was general agreement that updating of GC <111> was welcome.

A major concern was in the testing for similarity of dose-response curves. The revision appears to exclude hypothesis testing as described in the European and other pharmacopoeia, in WHO

guidelines and in much of the current bioassay literature, with a proposal to change to an equivalence testing approach.

A relevant question at this point was whether the revised chapter would be mandatory or a recommendation. It was reported at the USP Bioassay Workshop in August 2008, it had been indicated that the new chapters arising from the restructuring of GC <111> might be renumbered with four-digit chapter numbers, so that the proposed methods of analysis would be recommendations rather than being mandatory. Delegates at the BEBPA meeting commented that this could avoid some of the potential problems.

Within the Revision, there is inconsistency in requiring an equivalence testing approach for similarity. Line 56 states that “reasonable and sound alternatives may be employed”, but lines 203 and 206 state “do not use difference testing”. If “difference testing” refers to the F-test, this implies that the F-test is not “reasonable and sound”. Meeting attendees indicated a desire to select the most appropriate approach to similarity testing, as based upon the collection of data in their assay system. There was a strong feeling that, in this respect, proscriptive statements including phrases such as “do not use difference testing”, should not be included in the text.¹

If equivalence testing became mandatory, the following issues would need to be considered: This would compromise international harmonization, possibly requiring separate tests for different markets, with possibly inconsistent results

What would happen to testing of products already on the market?²

Equivalence testing requires the classification of statistically significant dissimilarity as being “trivial” or not. How and by whom is this classification to be made? One method recommended (line 205) is the use of historical data to determine the equivalence interval. However, line 301 notes that, in this case, “The equivalence interval specification will be driven solely by assay capability”.

The magnitude of dissimilarity in bioassay response would not necessarily reflect the importance of the consequences of the dissimilarity in a clinical application so any indication in bioassay data of dissimilarity in the product should not be dismissed as “trivial”

A Regulatory Authority testing a batch of product would not have historical data from its own laboratory to use in setting the equivalence limits

A manufacturer’s definition of equivalent might not agree with that of a Regulatory Authority

Another subject of concern to several delegates was the recommendation for use of an unweighted (geometric) mean for combination of potency estimates. In many cases only a small number of assays are performed and the use of a weighted mean may be desirable as the confidence limits on an unweighted mean will be wide. The use of a weighted combination assumes homogeneity, so this needs to be demonstrated. The proposed chapter states that an equivalence testing approach should be used (Section 6.2.2, point 1). The same concerns over the use of equivalence testing for assessing parallelism then arise in the testing for homogeneity.

Many delegates commented that a number of terms, such as “difference test” required definition. Other terms, such as “replicate” and “pseudoreplicate” were subsequently found in the draft glossary (*Pharmacopeial Forum* 2006;32(4):1359–1365). However, the fact that these terms were originally missed suggests that the structure of the draft glossary, possibly the divi-

sion into the various subsections, may not be easy for readers to consult. It was agreed that a comprehensive glossary was a necessity.

We hope that these comments will be helpful in the revision of USP GC <111>. Please note that this letter represents a summary of lively discussion, and while it cannot report all the comments and individual opinions expressed, we believe it reflects the major points made. It does not represent an official position taken by BEBPA or necessarily reflect the personal opinion of any member of the organizing committee

Sincerely yours,

Laureen E Little, PhD

Hans-Joachim Wallny, PhD

C Jane Robinson, PhD

The organizing committee of the BEBPA Inaugural Biological Assays Conference

The following additional comments were submitted by participants after viewing the draft of the letter:

1. The F test (and chi-square test) is performed on the differences between residuals of the curves (extra sum of squares). These residual differences themselves are a valid metric, firmly established in the statistical and pharmaceutical literature, and should be included in the chapter
2. The equivalence test can only be computed for linear systems. Nonlinear 4PL confidence regions can not be computed mathematically. Otherwise, this would need to be published in appropriate peer-reviewed statistical journals.