Product Quality Research Institute Evaluation of Cascade Impactor Profiles of Pharmaceutical Aerosols, Part 1: Background for a Statistical Method

Submitted: May 3, 2006; Accepted: October 13, 2006; Published: January 19, 2007

Wallace P. Adams,[1] David Christopher,[2] Douglas S. Lee,[3] Beth Morgan,[4] Ziqing Pan,[2] Gur Jai Pal Singh,[5] Yi Tsong,[6] and Svetlana Lyapustina[7]

[1]Office of Generic Drugs, Center for Drug Evaluation and Research, US Food and Drug Administration, Rockville, MD
[2]Statistics, Schering-Plough Research Institute, Kenilworth, NJ
[3]Nonclinical Statistics, Pfizer Global R&D, Groton, CT
[4]Manufacturing and Supply, GlaxoSmithKline, Zebulon, NC
[5]Former address: Office of Generic Drugs, Center for Drug Evaluation and Research, US Food and Drug Administration, Rockville, MD; current address: Watson Laboratories, Corona, CA
[6]Quantitative Methods Research Staff, Center for Drug Evaluation and Research, US Food and Drug Administration, Rockville, MD
[7]Pharmaceutical Practice Group, Drinker, Biddle and Reath LLP, Washington, DC

ABSTRACT

The purpose of this article is 2-fold: (1) to document in the public domain the considerations that led to the development of a regulatory statistical test for comparison of aerodynamic particle size distribution (APSD) of aerosolized drug formulations, which was proposed in a US Food and Drug Administration (FDA) draft guidance for industry; and (2) to explain the background and process for evaluation of that test through a working group involving scientists from the FDA, industry, academia, and the US Pharmacopeia, under the umbrella of the Product Quality Research Institute (PQRI). The article and the referenced additional statistical information posted on the PQRI Web site explain the reasoning and methods used in the development of the APSD test, which is one of the key tests required for demonstrating in vitro equivalence of orally inhaled and nasal aerosol drug products. The article also describes the process by which stakeholders with different perspectives have worked collaboratively to evaluate properties of the test by drawing on statistical models, historical and practical information, and scientific reasoning. Overall, this article provides background information to accompany the companion article's discussion of the study's methods and results.

KEYWORDS: Chi-square ratio, bioequivalence, cascade impactor, particle size distribution.

Corresponding Author: Svetlana Lyapustina, Drinker, Biddle and Reath LLP, 1301 K Street NW, Suite 900 East Tower, Washington, DC 20005-3317. Tel: (202) 230-5179; Fax: (202) 230-5379; E-mail: svetlana.lyapustina@dbr.com

INTRODUCTION

Drug deposition in the respiratory tract is influenced by the aerodynamic particle size distribution (APSD) of the active pharmaceutical ingredient from the drug product.[1-3] It is believed that, in general, aerosol particles greater than ~10 μm in aerodynamic diameter deposit primarily in the oropharynx and are swallowed rather than reaching the lungs. Smaller particles are thought to be deposited either centrally or peripherally in the lungs, depending on their size and the manner in which they are inhaled. The APSD measured in the laboratory by a cascade impaction method, for instance, is largely a characteristic of the delivery device, the formulation inside the device, and the APSD test procedure. Because of some plausible link between aerodynamic particle size and eventual deposition site within the respiratory tract, APSD may affect both the safety and the efficacy of orally inhaled and nasal drug products (OINDP).

When a manufacturer wishes to develop a generic version of a drug product, or when an innovator firm makes a change to a drug product and wishes to establish that the modified version is equivalent to the precursor (Reference) product, a critical issue is whether the new or modified (Test) product has an APSD sufficiently similar to that of the precursor product. This article presents a progress report of the Product Quality Research Institute (PQRI) working group that is analyzing the approach recommended by the 1999 US Food and Drug Administration (FDA) guidance[4] for investigating whether a Test product exhibits an APSD that is sufficiently comparable to that of the precursor product, with the focus on orally inhaled and nasal aerosols. The objective of this working group has been to recommend a robust method for assessing APSD equivalence, based on the guidance's approach or its modification or alternative. Even though the published guidance refers to only intranasal products, the focus of this working group from the beginning has been on all aerosols, including orally inhaled and nasal aerosols (but

excluding nasal sprays), which is explained by the agency's interest in applying a standard APSD profile comparisons test to all aerosol and dry powder inhaler OINDP.

The Methods section describes in detail the FDA rationale and objectives in developing the chi-square ratio test. This section also describes the process by which the proposed test could be thoroughly evaluated by scientists from the FDA, industry, academia, and the US Pharmacopeia (USP), using industry databases, extensive statistical modeling, scientific reasoning, and regulatory considerations.

The Results and Discussion section enumerates the features of a test that would make it ideal from both the scientific and regulatory perspectives. This list could be used as a benchmark for any new tests that might be proposed in the future for APSD profile comparisons.

Since the test under study was intended for practical application as a regulatory tool during review of drug applications, it is important to understand the boundaries of the test's design, to avoid any misunderstanding or misapplication. For this reason, the Additional Considerations section explains why the chi-square ratio is not designed to and should not be used for quality control within the Chemistry, Manufacturing, and Controls (CMC) context.

METHODS

FDA Test Development

In 1997, the FDA formed the OINDP Technical Committee (OINDP TC) to develop bioequivalence and product quality bioavailability guidances for OINDP intended for local action. For locally acting orally inhaled products formulated as metered-dose inhalers (MDIs) and dry powder inhalers (DPIs), the FDA's OINDP TC currently believes that both in vivo and in vitro tests are key components to establishing the equivalence of a generic to an innovator product. Furthermore, these tests can be envisioned to contribute to the characterization of individual product performance before and after the introduction of product changes likely to affect performance. The in vitro tests complement the in vivo tests, characterize the in vitro performance of the Test product relative to the Reference product, and provide additional assurance that the Test product is comparable to the Reference product as determined by the more variable and less sensitive in vivo tests. For equivalence of APSD, for which there is no established in vivo test, OINDP TC recommended an in vitro test based on cascade impactor (CI) measurements. The goal of the OINDP TC was to provide a test that (1) would be sensitive to potential differences between the Test product and the Reference product with respect to deposition at each individual deposition site in the impactor and accessories; (2) would be based on a single metric that incorporated differences on all sites to minimize the number of in vitro tests that must be met to claim equivalence; and (3) would be generally applicable to all inhalation products typically assessed in terms of APSD, rather than being drug product–specific.

A possible alternative to a test that considers deposition on each individual deposition site could be an equivalence test based on groups of deposition sites. For example, according to the draft CMC guidances for MDIs and DPIs,[5] for post-approval release and stability testing, but not for characterization of the drug product, drug deposition on individual sites may be grouped, with separate requirements placed on each of the groupings. However, because the APSD is drug product–specific, a general approach to establish the groupings is difficult to specify. Furthermore, combining 2 or more deposition sites within a single grouping (eg, a grouping composed of several adjacent impactor stages) would generally diminish sensitivity to differences between the Test product and the Reference product in deposition on individual sites in the impactor and accessories.

Another consideration of the FDA OINDP TC in selecting an equivalence test for APSD was the desire to have a single metric rather than multiple metrics, because for the Test product in vitro equivalence must also be demonstrated for other (non-APSD) tests, and the likelihood of failing by chance 1 or more tests in a battery of tests increases with the number of tests because of multiplicity. Therefore, a test based on independent comparisons of individual sites or groupings of sites was avoided.

Furthermore, no specific distribution was assumed for the APSD profile. If it had been established that an APSD could be described analytically by a mathematical function (eg, linear, lognormal, exponential), the parameters of that function could have served as the metric for comparison. However, no such function could be used in general to describe an APSD, and therefore a new metric and specifically designed test were sought.

The OINDP TC initiated work on developing a CI profile comparisons test in early 1998. Based on albuterol MDI data from an Andersen CI (USP Apparatus 1[6]), CI data were simulated at the FDA, with different mean deposition profiles and with specified levels of variability at each deposition site. Using the simulated data, the FDA conducted exploratory studies of the statistical performance and appropriateness for APSD profile comparisons of 4 ratio tests for equivalence testing, and it was determined that the chi-square ratio test appeared most promising. These 4 ratio tests were based on the statistics of the ratio of (1) chi-squares, (2) mean square differences (Cramer–Von Mises statistic) of cumulative percentages, (3) mean absolute differences of cumulative percentages, and (4) similarity factors (f2) of cumulative percentages. Each of these original statistics provides a measure of the distance between 2 profiles. The ratios were constructed

with the Test-to-Reference distance in the numerator, and the Reference-to-Reference distance in the denominator. More complex tests might potentially be created, but one of the objectives was to have a relatively uncomplicated test to be implemented in practice.

An APSD CI profile is composed of deposition (mass) data on multiple sites in the impactor. If we had counts of numbers of particles on each stage instead of mass, we would have an ordered multinomial distribution. The chi-square statistic is often used to test the significance of differences between such multinomial distributions, and its properties have been well studied for that use.[7] For application to comparison of CI profiles, we adopt the form of the chi-square statistic to provide a measure of the distance between 2 profiles. This statistic is calculated as the sum of the squared differences in deposition at each impactor site (differences between the 2 profiles), scaled by the average deposition on that site (averaged between the 2 profiles). In developing a test for APSD profile comparisons, the FDA OINDP TC used this statistic to construct a new metric, namely a ratio of the chi-square distance of Test-to-Reference to the distance of Reference-to-Reference. (The corresponding formulas are provided in the 1999 Draft Guidance[8] and PQRI documents.[9]) This was done to have the test react to the differences in variation between the Test and Reference profiles in addition to the distance between the 2 profiles. To determine a chi-square ratio value (a so-called critical value) that would allow equivalent profiles to be distinguished from profiles that failed to establish equivalence, the FDA OINDP TC used simulation studies of Test and Reference profiles with identical mean depositions but different variabilities at each site, and of Test and Reference profiles with different mean depositions and fixed variabilities at each site. One critical value from such preliminary studies was reported at a 2000 meeting.[10] (A detailed theoretical description of this and the other 3 tests considered by the FDA OINDP TC is available from the PQRI Web site.[9]) These analyses revealed that the ratio of chi-squares test performed better than the other tests investigated by the FDA OINDP TC.[11] Therefore, the ratio of chi-squares test was proposed in the June 1999 "Draft Guidance for Industry: Bioavailability and Bioequivalence Studies for Nasal Aerosols and Nasal Sprays for Local Action"[4] and later was posted as a stand-alone appendix[8] to the updated draft of the guidance.[12]

Assessment of Test Performance: Mandate of PQRI Working Group

In 1999, the FDA and several trade and professional associations established PQRI[13] as a collaborative process between the Center for Drug Evaluation and Research (CDER)/FDA, industry, academia, and the USP, to conduct research that explores scientific bases for regulatory policy related to drug product quality. PQRI is governed by a board of directors and a steering committee. The steering committee oversees TCs that currently focus on research projects in 4 disciplines: drug product, drug substance, biopharmaceutics, and manufacturing. Each TC defines research objectives for its working groups, while the working groups prepare work plans detailing the specifics of the work to be accomplished. After approval of the work plan by the PQRI senior committees, a working group implements the plan and usually publicizes its findings through scientific journals and conferences as the work progresses. Upon completion of the work, a report with recommendations for changes to existing or draft guidances is prepared and submitted to the FDA/CDER for consideration. CDER reviews the submission and responds to PQRI, either accepting the recommendation or explaining why it cannot be accepted as given and what additional information is needed.[14]

In late 2001, a PQRI Working Group was formed under the Drug Product TC (an Aerodynamic Particle Size Distribution (APSD) Profile Comparisons Working Group), which included representation from the above-mentioned 4 sectors. As detailed in its 2002 workplan,[15] the working group set as its objective the examination of the chi-square-based method proposed by the FDA and, if needed, revision of this test to develop a more widely applicable method for testing the equivalence of APSD profiles. The working group is focusing on MDI and DPI data from the Andersen CI, either as USP Apparatus 1 or as USP Apparatus 3, but is also interested in extending the approach to other CIs, in particular the next-generation pharmaceutical impactor, which may become the impactor of choice in the future and is now included in USP chapter <601> as Apparatus 5 and Apparatus 6.

The status of this project and issues under consideration by the working group have been presented at several public meetings.[16-18] The companion article presents the work and findings to date and outlines the areas to be addressed by the working group in the near future to determine appropriate recommendations to the FDA on this topic.

Results and Discussion

Ideal Profile Comparisons Test

In the course of its investigations, the working group had to consider a number of questions, such as what constitutes a good test and how sensitive, discriminating, and consistent a test should be.

These discussions resulted in a consensus list of desirable characteristics, which are presented below. They help set

the stage for the studies performed by the working group, which are described in the companion article.

- The test and critical value should be independent of a particular impactor type or impactor configuration.
- The test should be applicable to a broad range of orally inhaled and pressurized nasal products.
- It is desirable that the test be applicable to a broad range of realistic profiles; however, it is not necessary that the test be applicable to all theoretically conceivable profiles.
- The test should reward the Test product when it has lower variability than the Reference product, based on the FDA goals for the test.
- The test statistic should have a minimum for an ideal Test product. An ideal Test product has the same mean as the Reference product and zero variance; thus, the lower variability in the Test product is rewarded.
- An ideal test would not require any distributional assumptions about deposition on individual stages.
- It is desirable but not essential that the test statistic itself (eg, the 95th percentile of the distribution of the mean of the chi-square ratios), or at least the mean of the chi-square ratios, follow a known distribution, such as normal, chi-square, or F. (The approach of analyzing a test's performance by looking, in particular, at the distribution of means or other selected percentiles has been explained in detail elsewhere.[19])
- The test should be sufficiently sensitive without being overly sensitive because from a practical point of view, exaggerated sensitivity might limit the usefulness of the test as a decision-making tool. For example, a statistical test may be able to detect small differences in the means without such differences being of any practical significance for the given regulatory task. On the other hand, a test may not be able to extract relevant information in a consistent way because of a low signal-to-noise ratio, or high variability in the data relative to the test's capabilities. With either extreme, the ability to make correct decisions consistently would be hampered. Ideally, a test should be sensitive and powerful enough to react to the differences that are considered important, but only to those that are important. For example, the working group had to consider (1) whether sensitivity to the change in the mean or sensitivity to the change in variability is more important, or whether they are equally important; and (2) whether the ideal test's sensitivity should be equal for all sites or should be greater for some sites than for others (eg, with greater sensitivity to sites with higher deposition compared with sites with lower deposition, or with greater sensitivity to fine particle deposition sites).

The above list details the ideal properties of a profile comparisons test. However, the objective of the working group has been to determine whether the chi-square ratio test can make consistently correct decisions about Test and Reference profile comparisons even if the test does not have all of the ideal properties.

Additional Considerations

APSD Profile Comparisons Test Is Not Appropriate for Quality Control Purposes

An APSD profile comparisons test, as described in the FDA published guidance and discussed in this article, is intended to compare a Test product (eg, a generic product, or an innovator product after certain changes in device components, formulation, or supplier) to a Reference product. The developed test metric is a *ratio* of the 2 chi-square statistics characterizing both products. Thus, it would be nonsensical to speak about a predetermined "chi-square ratio specification" for a given, isolated profile. Moreover, to obtain meaningful results from the chi-square ratio calculation, both of the products must be tested within the same set of experiments and at the same time. This precludes use of the profile comparison test throughout stability studies.

In summary, the chi-square ratio test was not designed as a quality control test. Its properties in that context are unknown. Thus, its use for quality control purposes would be inappropriate and incompatible with its design and intentions.

Zero Deposition Sites

Zero (or below the limit of detection) deposition on individual sites may pose a challenge in some implementations of the algorithm,[20] but this situation should be manageable as long as the 2 Reference profiles used in the calculation of the chi-square ratio are not identical.[21] From the formulas for the chi-square ratio, it is clear that theoretically, the mathematical problem arises either when all of the Reference sites have zero deposition, or when both of the Reference profiles used in the calculation are exactly the same. The first of these situations is unlikely to occur in practice, and the second has been eliminated by the definition of the algorithm (ie, 2 Reference profiles, R1 and R2, are selected such that $R1 \neq R2$).

CONCLUSION

An objective statistical test to determine the equivalence of CI profiles of pharmaceutical aerosols would be a helpful tool for both industry and regulators. A draft FDA guidance for industry recommended a particular test that has been further evaluated through a PQRI working group. This article, which explains the working group's background and mandate, is the first in a series documenting the working group's investigations and findings. The interim results are presented in the second article of this series.

ACKNOWLEDGMENTS

This article presents the professional opinions of the authors and does not necessarily represent the views or policies of the US Food and Drug Administration or any other institution with which the authors are affiliated. The working group would like to thank the PQRI and all its member organizations for encouraging and supporting this work.

The authors also thank the AAPS PharmSciTech Editor-in-Chief Pat DeLuca and the anonymous referees for helpful comments and suggestions made during review of the manuscript.

On behalf of the PQRI APSD Profile Comparisons Working Group, additional support and information was supplied by the following contributors: Office of Generic Drugs, CDER, FDA, Rockville, MD; Statistics, Schering-Plough Research Institute, Kenilworth, NJ; Nonclinical Statistics, Pfizer Global R&D, Groton, CT; Manufacturing and Supply, GlaxoSmith-Kline, Zebulon, NC; Statistics, Schering-Plough Research Institute, Kenilworth, NJ; During the PQRI work reported here: Office of Generic Drugs, CDER, FDA. Current address: Watson Laboratories, Inc. Corona, California; Quantitative Methods Research Staff, CDER FDA, Rockville, MD; Pharmaceutical Practice Group, GCD, Washington, DC; The PQRI APSD Profile Comparisons Working Group is chaired by David Christopher (Schering-Plough Research Institute). Members of the Working Group (in alphabetical order) are: Wallace P. Adams (FDA), Anton Amann (ACN Pharma), Craig M. Bertha (FDA), Peter R. Byron (Virginia Commonwealth University), William H. Doub (FDA), Craig Dunbar (Alkermes), Walter W. Hauck (Thomas Jefferson University, current address: USP), Douglas S. Lee (Pfizer), Richard Lostritto (FDA), Svetlana Lyapustina (Drinker, Biddle and Reath LLP), Jolyon P. Mitchell (Trudell Medical Group), Beth Morgan (GlaxoSmithKline), Steve Nichols (Aventis), Ziqing Pan (Schering-Plough Research Institute), Gur Jai Pal Singh (FDA), Terrence Tougas (Boehringer Ingelheim), Yi Tsong (FDA), Ronald K. Wolff (Nektar), Bruce Wyka (Schering-Plough Research Institute), and the Working Group's liaisons to the PQRI Drug Product Technical Committee Michael Golden (GlaxoSmithKline) and Guirag Poochikian (FDA).

REFERENCES

1. Brain JD, Blanchard JD. Mechanisms of particle deposition and clearance. In: Morén F, Dolovich MB, Newhouse MT, Newman SP, eds. *Aerosols in Medicine.* New York, NY: Elsevier; 1993:117–156.

2. Gonda I. Targeting by deposition. In: Hickey AJ, ed. *Pharmaceutical Inhalation Aerosol Technology.* New York, NY: Marcel Dekker; 1992:61–82.

3. Hinds WC. *Aerosol Technology.* New York, NY: John Wiley & Sons; 1999:233–259.

4. FDA. CDER. Draft guidance for industry: bioavailability and bioequivalence studies for nasal aerosols and nasal sprays for local action. 1999; Available at: http://www.fda.gov/cder/guidance/2070DFT.pdf. Accessed April 17, 2006.

5. FDA. CDER. Draft guidance for industry: metered dose inhaler (MDI) and dry powder inhaler (DPI) drug products—chemistry, manufacturing, and controls documentation. 1998; Available at: http://www.fda.gov/cder/guidance/2180dft.pdf. Accessed April 17, 2006.

6. USP. Chapter <601> aerosols, nasal sprays, metered dose inhalers, and dry powder inhalers (in-process revision). *Pharm Forum.* 2003;29:1176.

7. Johnson NL, Kotz S, Balakrishnan N. *Continuous Univariate Distributions*. New York, NY: John Wiley; 1994:415–493.

8. FDA. CDER. Statistical information from the June 1999 draft guidance and statistical information for *in vitro* bioequivalence data. 1999; Available at: http://www.fda.gov/cder/guidance/5383stats.pdf. Accessed April 17, 2006.

9. Tsong Y. Statistical comparison of particle size distribution profiles. 2004; Available at: http://pqri.org/commworking/minutes/pdfs/dptc/psdpcwg/Addl/DC01-475116-v2-Yi_Tsong_Statistical_Archive_PQRI_Profile_Comparisons.DOC. Accessed April 20, 2006.

10. Tsong Y. Profile analysis of cascade impactor data: proposed FDA approach. Presentation to OINDP Subcommittee of the Advisory Committee for Pharmaceutical Science. Rockville, MD. April 26, 2000. Available at: http://www.fda.gov/ohrms/dockets/ac/00/slides/3609s1e.ppt. Accessed April 17, 2006.

11. Tsong Y. Comparative statistics for assessing in vitro equivalence based on cascade impactor distribution profiles. Presentation at: AAPS/FDA/USP Workshop Regulatory Issues Related to Drug Products for Oral Inhalation and Nasal Delivery; June 3-4, 1999; Washington, DC.

12. FDA. CDER. Draft guidance for industry: bioavailability and bioequivalence studies for nasal aerosols and nasal sprays for local action. 2003; Available at: http://www.fda.gov/cder/guidance/5383DFT.pdf. Accessed April 17, 2006.

13. PQRI Homepage Web site. Available at: http://www.pqri.org. Accessed August 17, 2006.

14. Massa T. Overview of PQRI's structure, objectives and process. Presentation at: PQRI Conference Good Regulation Through Good Science; August 5-6, 2003; Arlington, VA.

15. PQRI PSD Profile Comparisons Working Group. Work plan for the investigation of an optimized chi-square method for comparing particle size distribution profiles obtained by cascade impactors with specific reference to equivalence testing of orally inhaled and pressurized nasal drug products. 2002; Available at: http://www.pqri.org/commworking/minutes/pdfs/dptc/psdpcwg/workplan02.pdf. Accessed August 18, 2006.

16. Christopher D, Lee D. Report on particle size distribution profile comparisons working group. Presentation at: PQRI Conference Good Regulation Through Good Science; August 5-6, 2003; Arlington, VA.

17. Christopher D. Statistical assessment of the proposed chi-square method for equivalence of particle size distribution profiles of cascade impactor data. Presentation at: The 7th Conference on Nasal and Pulmonary Drug Delivery; September 17, 2003; Barcelona, Spain.

18. Lee D. Searching for the holy grail of a single PSD profile comparator: technical challenges with the chi-squared ratio approach. In: Dalby RN, Byron PR, Peart J, Suman JD, Farr SJ, eds. *Respiratory Drug Delivery IX: Biological, Pharmaceutical, Clinical, and Regulatory Issues Relating to Optimized Drug Delivery by Aerosol.* vol. 1. River

Grove, IL: Davis Healthcare International Publishing; 2004: 161–169.

19. Christopher D, Pan Z, Lyapustina S. Aerodynamic particle size distribution profile comparisons—considerations for assessing statistical properties of profile comparison test. *Am Pharm Rev.* 2005;8:68–72.

20. Byron PR. *In vitro* bioequivalence of particle size distributions: a commentary on FDA's draft guidance. In: Dalby RN, Byron PR, Peart J, Suman JD, Farr SJ, eds. *Respiratory Drug Delivery VIII: Biological, Pharmaceutical, Clinical, and Regulatory Issues Relating to Optimized Drug Delivery by Aerosol.* vol. 1. River Grove, IL: Davis Healthcare International Publishing; 2002:145–154.

21. PQRI PSD Profile Comparisons Working Group. Minutes of the teleconference on 7 July 2004:3. Available at: http://www.pqri.org/commworking/minutes/pdfs/dptc/psdpcwg/070704min.pdf. Accessed April 17, 2006.