COMMENTS ON DRAFT GUIDANCE FOR INDUSTRY: BIOAVAILABILITY AND BIOEQUIVALENCE STUDIES FOR NASAL AEROSOLS AND NASAL SPRAYS FOR LOCAL ACTION (2ND DRAFT, APRIL 2003) [DOCKET NO. 99D-1738]

We agree with FDAs recognition of the challenges/difficulties of assessing bioavailability and bioequivalence of nasal sprays/aerosols, particularly of corticosteroids intended for local action. Thus, in general, we support generation of this guidance and public definition of rigorous scientific standards against which interchangeability between Test and Reference products may be established. Our overall comments are shown below, followed by our specific comments. Specific comments are organized under the same section headings as used in the draft guidance and cross-referenced by line number. All section headings are included.

OVERALL COMMENTS

1. The statistical standards for determining bioequivalence for both *in vitro* and *in vivo* studies were not provided for public comment and thus this guidance is incomplete.

FDA acceptance ranges for the assessment of equivalence (*in vitro* and *in vivo* studies, pharmacokinetics (PK) and pharmacodynamics (PD)) should be stated and justified in the document. Appendices were not made available to industry during this comment period. A full understanding of the risk to the public (type I error) and risk to sponsors (type II error) may not be made in the absence of the statistical appendices. Thus we suggest that, as a minimum, an additional period of public comment, and probably a public meeting, will be required following publication of these statistical appendices before this guidance may be finalized. The need to establish standards for therapeutic equivalence of intranasal products provides a good opportunity to review the available scientific information in a public forum.

The use of the population bioequivalence approach has been shown to be insensitive to mean changes in formulation due to scaling to reference product variability. The FDA should carefully consider the choice of acceptance range if this approach is to be applied to *in vitro* testing and ensure that the choice of acceptance range for such assessment of equivalence protects public health.

It is assumed that FDA intends to apply average bioequivalence testing techniques to the *in vivo* studies (PK and PD). This should be clarified in the guidance.

2. We strongly support use of PK studies to assess systemic absorption and these should be required even if it is not possible to obtain a full PK profile. The most sensitive methodology should be selected.

The PK and PD studies described in the draft guidance have very different objectives and are not interchangeable. We do not agree that a 6-week HPA axis study in adults would substitute for PK measures.

3. We do not agree that a PD study in adults would support extrapolation of a determination of Bioequivalence (BE) to children. If it proves impossible to conduct a rigorous PK study in healthy adult volunteers, a PD study in adults should be supplemented by a PD study in children to assure BE in this patient subgroup.

Effects on urinary free cortisol or serum cortisols may not be a sufficiently sensitive measure of HPA axis function for the purpose of making a determination of bioequivalence or interchangeability, as evidenced by the FDAs class labeling for inclusion in the Pediatric Use section of package inserts for intranasal corticosteroids i.e., "This effect [on growth] has been observed in the absence of laboratory evidence of HPA axis suppression, suggesting that growth velocity is a more sensitive indicator of systemic corticosteroid exposure in pediatric patients than some commonly used tests of HPA axis function". If it proves impossible to conduct a sensitive, rigorous PK study in adults, a PD study in adults would need to be supplemented by a PD study in children, using a sensitive measure such as growth velocity.

4. More discussion is required around the most appropriate model of local efficacy. A perennial allergic rhinitis (PAR) clinical study should be the recommended indication for a test of equivalent efficacy.

Although seasonal and perennial allergy patients experience the common symptoms of sneezing, itching, rhinorrhea and nasal congestion, studies have shown that the perennial sufferer, both allergic and nonallergic, experiences more severe and sustained nasal congestion. The proposed two-week SAR efficacy study may provide confidence that the Reference and Test formulations are statistically superior to placebo for efficacy endpoints, but the study would provide little confidence that the two products would perform similarly if one product were substituted for another; the study is not powered to demonstrate that the Reference and Test product would provide similar responses to treatment and no standards for determining bioequivalence are provided. We do not understand if this study is intended to perform a "confirmatory" role or a "pivotal" role, as referenced in Dr. Meyer's presentation at the July 17, 2001 OINDP Subcommittee meeting. We maintain that onset of effect should be included as a parameter of interest in a traditional treatment study in PAR patients, especially since dose-response relationships are difficult to demonstrate for TNSS efficacy parameters.

5. The same standards should apply to ANDAs and NDAs.

We support application of the same standards to ANDAs and NDAs, as product quality, safety and efficacy considerations are independent of the regulatory mechanism for approval. The public interest will be served by the adoption of scientifically appropriate equivalence criteria that will ensure that approved Test (generic) formulations of these products are interchangeable with the Reference listed drugs that have been proven safe and effective and that changes between development and to-be-marketed products are validated.

However the requirements for Bioavailability (BA), BE and comparability are not interchangeable, nor are they clearly explained in this guidance.

6. The guidance needs to be very clear on the standards for *in vivo* and *in vitro* tests associated with making a determination of therapeutic equivalence and substitutability; the *in vivo* and *in vitro* studies appropriate to determination of BA of a development product, or the assessment of BE associated with validating changes between a clinical trial and to-be-marketed product.

The nomenclature describing comparisons in each section should be consistent with past conventions. The term "bioequivalency" is used throughout. We recommend using "pharmaceutically equivalent" for *in vitro* testing, "bioequivalent" for PK and PD testing, and "clinically equivalent" for local efficacy testing. With specific regard to *in vitro* tests, the terms "equivalence", "comparability" and "substitutability" should be explicitly defined and differences in supportive data requirements clearly explained.

7. The relevance and role of the *in vitro* tests in the assessment of BA and BE warrants further discussion.

We recommend that the standards are limited to those with proven relevance and discriminatory power. We agree with the statement in the draft guidance that acknowledges that comparable *in vitro* performance cannot be extrapolated to *in vivo* BE at this time simply because the clinical models currently available are not sufficiently sensitive to discriminate between Test and Reference products.

8. The *in vitro* tests described in this guidance should be consistent with those described in the companion CMC guidance.

The final guidance "Chemistry, Manufacturing and Controls (CMC) Guidances For Industry: Nasal Spray and Inhalation Solution, Suspension and Spray Drug Products" and the current draft BA/BE guidance are significantly clearer than the previous drafts but the relationship between the two guidances should be clearly stated. It is our understanding that NDA and ANDA submissions would need to conduct all tests described in the CMC guidance during product development, but assessment of the interchangeability between Test and Reference would depend only on the results of tests described in the BA/BE guidance and compliance with BE standards that have been identified *a priori*.

DETAILED COMMENTS, ANNOTATED TO EACH SECTION OF THE DRAFT GUIDANCE

I. INTRODUCTION

Line 23: We welcome the fact that this guidance applies equally to Nasal Aerosols (MDIs) and to Nasal Sprays (aqueous), and to NDA and ANDAs.

Line 24: The scope of the guidance is Nasal Sprays, however the Division of Pulmonary and Allergy Drug Products (DPADP) applies the concepts of this guidance to products for oral inhalation. This has been specifically noted in meetings and written communications. It should be clearly stated that the DPADP may: apply these concepts to other products, or specific dosage form guidance should be published, or guidance provided on the Agency's intention with respect to development of parallel guidance on oral inhalation products.

It is not possible for us to determine the extent of representation from the DPADP on the committees listed in footnote One, but we assume that they have been fully consulted and will participate in the rest of the standard setting process.

Lines 39-40: The comment period should remain open until these appendices, which are an integral component of the guidance, are available, allowing for an adequate review and comment period.

II. BACKGROUND

Line 62: Either delete specific reference to Item 6 or update this to include reference to section 5.3.1 of the Common Technical Document format NDA.

Line 63: All the *in vitro* tests defined focus on product performance, however it is not clear that they all define how well the drug substance is released from the product. The parentheses should be deleted from this sentence.

Line 67: It is not clear how this guidance should be used with other more general CMC guidance. Should a comparison with other CMC tests not defined here be part of the equivalency assessment? If a product passes the *in vitro* criteria, but the results of other CMC tests are not aligned would the product be considered BE (i.e., should a product have an equivalent Impurity profile, difference's in weight loss between Test and Reference?). These may represent products of differing product quality, however equivalent based on this guidance. Acceptable differences for all CMC tests defined in CMC guidance should be specified.

Line 77 (footnote 4): We are not sure that this reference to the draft MDI/DPI guidance is appropriate. Perhaps this should be a reference to the aforementioned Nasal Spray guidance.

A. BA and BE Data

Lines 93-95: We agree with these criteria but observe that they are not provided with this guidance and thus the document is substantially incomplete.

Lines 102-104: We suggest revision to this sentence to "While a drug administered nasally and intended for local action **may** produce systemic activity, plasma levels do not in general reflect **clinical efficacy or** the amount of drug reaching nasal sites of action".

1. Local delivery BA/BE concepts

Line 115: Droplet size distribution and deposition pattern from a pump-style nasal spray are not necessarily dependent on the drug substance. We recommend that this sentence be amended as follows: "....produces droplet or drug particle sizes and distribution patterns within the nose that are dependent upon the formulation and device (container-closure system), and, in the case of nasal aerosols, the drug substance characteristics."

Lines 123-133: According to the guidance document, FDA recognizes that there may be limitations in understanding the clinical relevance of *in vitro* tests, however it has been our experience that Agency expectations for these tests have escalated. We do agree that "the clinical relevance of these tests, or the magnitude of the differences in the tests, is not established" and we strongly agree with the statement in lines 132-133 that "clinical studies can unequivocally establish the effectiveness of the drug product".

Line 135: Insert the following sub-heading for ease of reference "BA and BE Assessment of Solution Formulations".

Line 144: Insert the following sub-heading for ease of reference "BA and BE Assessment of Suspension Formulations".

Lines 153-156: Bioavailability does not necessarily reflect availability at the site of action. It may help describe the degree of nasal deposition but it should be noted that where oral absorption is the major source of systemic exposure the data are unlikely to be informative about topical exposure in the nose.

2. Systemic exposure and systemic absorption BA/BE concepts

Lines 159-162: We agree that it is more desirable to measure pharmacokinetics (PK) when possible and to use pharmacodynamic (PD) measures when this is not possible. It should be clearly stated that the clinical endpoints in the local efficacy study and the safety related endpoints in the PK/PD studies are not directly related. PK may be used as a tool to measure product performance, but not as a surrogate for efficacy of these products.

Lines 163-169: It is not clear what the threshold is for determining whether PK studies are feasible. In order to obtain measurable levels of systemic drug, it may be necessary to administer greater than therapeutic doses. Generation of data at higher than therapeutic doses would provide valuable information for safety purposes and may be justifiable in order to generate PK profile data. However, we agree that great care must be used when using supra-therapeutic doses because of the risk of saturating the system. If most of the dose is swallowed or runs out the nose, the data are of limited value. There should be some rationale for using a supratherapeutic dose besides inability to obtain measurable concentrations. This could be accomplished in a pilot study by examining multiples of the therapeutic dose to demonstrate a dose-related increase in exposure. That would provide an indication that the system is not being saturated and the dose chosen is sufficiently sensitive. Additionally, information on the relative contributions to systemic exposure from nasal and oral absorption would be a guide to the relevance of PK data to topical efficacy (Reference 1).

Line 170: We welcome the inclusion of a Decision Tree but are unable to provide comments at this time since this was not circulated with the guidance. We recommend that it include a determination of Q1, Q2, confirmation of matching container closure systems, and clarification of the role of PD systemic exposure studies, with distinct separation of the standards for assessing bioavailability and determination of "therapeutic equivalence".

B. CMC Tests and *In vitro* BA Tests (Noncomparative) versus BE Tests (Comparative)

Lines 176-180: The inference that CMC release tests do not "focus on the release of drug substance from the drug product" is a somewhat sweeping statement. For example, the Nasal Spray CMC guidance states "Comprehensive and well-defined *in vitro* performance characteristics should be established before initiating critical clinical or bioequivalence studies" in the specifications section.

Lines 180-184: We agree with the stated distinctions between BA and BE. We agree that BE limits should be established *a priori* and believe that they should be standardized across the industry for a single product, based on the variability of the Reference product. A determination of substitutability of each Test product with the Reference product should be dependent on compliance with the same BE limits and these limits should be transparent. Unfortunately, no bioequivalence standards were provided in this guidance.

III. FORMULATION AND CONTAINER AND CLOSURE SYSTEM

A. Formulation

Line 197: There is no discussion of the importance of assuring equivalent quality of the drug substance and excipients. This is important because of the potential impact of a differing impurity profile on the overall quality of the finished product, and would be consistent with the draft Guidance for Industry: ANDAs: Impurities in Drug Products (December 1998).

Line 199: For suspension products the PSD of the drug substance should be equivalent to the reference product, not just comparable. The Agency should define comparable.

Line 205: Define identical. It is not realistic that any 2 batches of drug substance will have identical PSD. However, the PSD acceptance limits applied to the T and R product and approved in the application should be identical.

B. Container and Closure System

Lines 209-212: We agree that all these components are an integral part of the product. There are some instances where the protective packaging might have no contact with the product and would not affect product performance where it would be appropriate to recommend that sponsors initiate a dialogue with the Agency.

Line 217: We agree that use of the same brand and model of device is the best way to assure equivalence. The scientific support for this is provided in lines 219-225. If this is not feasible, there are likely to be some differences in the materials of construction of the container closure system between an ANDA and the Reference listed drug and it would be necessary to ensure that the extractives profiles of closure components did not have an adverse effect on the quality of the final product, as well as the physical characteristics of the components. The guidelines should place an appropriate and critical emphasis upon sponsors working with pump and actuator suppliers to understand the moulding/assembly parameters that need to be controlled in order to meet the stringent expectations of the agency.

Line 222: It is our understanding that the orifice geometry and swirl chamber design refer to the actuator design rather than the pump thus this sentence could be clarified as follows "... including the precompression mechanism and actuator design, (including specific geometry of the orifice (Kublic and Vidgren 1998), and the design of the swirl chamber)".

Line 224: We suggest that the Agency clarify that the external dimensions of the Test actuator are expected to ensure comparable depth of insertion to the Reference actuator in order to make a determination of interchangeability and would not apply to circumstances where confirmation of BA is intended to justify a novel design with greater patient acceptability. Where the document specifies that all the device parts should be identical, it does not mention the outer casing. In theory this might influence the orientation of the tip of the device but also have an impact on compliance and thus

efficacy. We suggest that this would be essential to a determination of BE but would not apply to BA assessments conducted to establish comparability during new drug development.

Line 225-229: To minimize confusion to the patient and Healthcare Practitioner (similar to the requirement for dose proportionality for multiple strengths) the number of shots to prime the T and R product should be identical. The number of shots to prime the pack affects the amount of overfill included and ability for the patient to achieve the labelled number of sprays. However, this requirement would not apply to a line extension in a novel intranasal device. We suggest that the text be amended as follows: "For generic products, the Test product is expected to attain prime within the labeled number of actuations for the Reference product."

IV. DOCUMENTATION OF BA AND BE

A. NDAs

Lines 236-37: This implies a new NDA filing requirement to establish a "baseline" and all the *in vitro* tests defined here are not fully described in the June 2002 final Nasal Spray CMC guidance.

Line 240-244: We agree with this advice.

B. ANDAs

Lines 250-264: We agree with the proposed conventions for product equivalency. We suggest that some reference to equivalent quality of components be included.

1. Solution formulations

2. Suspension formulations with PK systemic exposure data

Lines 280-291: We strongly agree that both *in vitro* and *in vivo* studies are required to document BE for suspension formulation products intended for local action.

Lines 283-285: A PK study is valuable to define *in vivo* product performance even if plasma concentrations cannot be measured at all time points.

3. Suspension formulations without PK systemic exposure data

C. Postapproval Change

V. IN VITRO STUDIES

A. Batches and Drug Product Sample Collection

1. NDAs

Lines 318-322: We strongly agree that the three batches selected for *in vitro* BA studies should be as representative of the to-be-marketed product as possible.

Line 328: The suggestion to test 3 batches (i.e., clinical, stability, production) at the same time might be difficult given that the manufacture dates can be far apart (e.g., years). The recommendation should be limited to 3 batches tested, with a suggestion to test them at the same time, if possible.

Line 335: The suggestion to provide all completed batch records is excessive, especially if the early batches are manufactured with a pilot scale system that is not representative of the to-be-marketed batches. Provision of representative batch records from a production batch is more appropriate. The guidance could suggest provision of a summary between batches to supplement production records.

2. ANDAs

Lines 349-352 and 356-361: We strongly agree that BE testing should be conducted on three representative batches of Test and Reference product.

Line 350: Consideration should be given to using batches of different age to evaluate *in vitro* BE to ensure that the Test product performance is equivalent to the Reference product throughout the approved shelf life. These tests should be completed on batches stored under the labeled conditions for the full duration of the Reference product shelf life. We suggest that the phrase "or more" is removed from the sentence and that a requirement be added that BE comparisons be conducted throughout (2 or more points) the Reference product shelf life.

B. Tests and Metrics

Line 378: The recommendation to use laser diffraction for PSD of droplets for Nasal Aerosols may be very difficult to validate and interpret due to high variability.

Lines 390 and 402: Duplication. Move reference to footnote 11 to line 390 and delete redundant sentence in line 402/3.

Lines 400-401: We strongly agree with all references to the importance of eliminating potential sources of bias during collection of these data. This suggests that such studies should have a protocol containing pre-defined criteria for randomization, batch selection, variability, replicate analyses for each test and a determination of BE. We suggest that this be specifically stated and that the standards for determination of BE should be established *a priori*. However we recognize that blinding the analyst to the sample identity will significantly complicate the analysis. We suggest that it is made clear that this should only apply to BE determinations where all container-closure components etc., are intended to be interchangeable, not to BA assessments intended to validate changes made during development of new products.

Line 406: We would not normally report experiments rejected due to assignable causes (e.g. instrument failure). We request clarification of the requirement to see data that have been shown to be meaningless by laboratory investigation. We suggest revision of this sentence as follows "…replaced during *in vitro* analyses and failure to use the specific actuations required by the protocol. The original and reanalyzed data…." This point is especially relevant if the assignable cause is based on the results of established, well-controlled procedures for conducting out-of-specification investigations.

Lines 413 and 423: In line 413 the agency requests "all raw data" while in line 423 20% of the total observations are required to be submitted. This requirement is extremely onerous and will lead to enormous CMC sections with no added value. We suggest that the requirement for data is clarified and limited to individual results and representative output from instrumentation.

Lines 423-426: We suggest that the Agency state whether their preference is for electronic images or the electronic raw data files themselves.

1. Single actuation content (SAC) through container life

Lines 430-451: "Single actuation content through container life" is a new test and should be added to the CMC guidance. Single actuation content can be influenced by the physical age of the suspension; therefore this test should be applied throughout the shelf life of the Reference and Test product at equivalent ages.

Line 443: It is not clear what advantage a stability-indicating chemical assay provides in evaluating this criterion. Assuming equivalent suspension formulations/content (based on successful assay, spray content uniformity results), shot (spray) weights would appear to be equally effective in demonstrating comparable single actuation content. Alternative methods should be available to sponsors.

Line 446: Consistent with the CMC guidance (III.F.1.g), this test should represent the usual or minimum dose described in the product labeling, i.e., if the dose is one spray per nostril, a single dose is two sprays, one in each nostril. It is our opinion that emitted dose comparisons between Test and Reference products should always be made based on the labeled number of actuations. SAC could lead to erroneous conclusions.

Line 448: Consistent with the CMC guidance (III.F.1.g), this test should measure the combined BOU and EOU. The requirement for MOU should be removed.

Line 459: Clarify 'lagering period'. Should this mean 'quarantine period'?

Lines 459-467: It is not clear why these comments on priming appear in the BA/BE guidance, rather than the companion CMC guidance.

Line 467: Reference to batch release testing with respect to the priming SAC study is confusing as the SAC test and batch release tests are not typically conducted on the same inhalers or using the same method.

2. Droplet size distribution by laser diffraction

Lines 471-522: It is not clear how the data should be evaluated. The recommendations for the test and data presentation seem overly prescriptive.

a. Nasal sprays

Line 482: Can the Agency clarify whether other distribution parameters can be justified, e.g., D10, D50, D75? It is not clear whether D90 and "span" are requirements.

Lines 495-503: We agree that measurement of the stable part of spray removes variability due to formation of the droplet at initial and dispersion of plume at the conclusion.

Line 512: The variability of laser diffraction will make single spray DSD results meaningless, additional detail is needed in the guidance for sponsors to understand why this is required.

Line 516: Can the Agency confirm that assessment of droplet size distribution within the fully developed phase will be acceptable for routine quality control testing of nasal spray products? An explanation of the rationale for conducting measurements of BOU and EOU and measurement at two distances would be helpful. In our experience, adequate information regarding product quality and performance can be obtained simply from making measurements at one distance. While it can be argued that such information can only be improved by making measurements at two distances, it is questionable whether this level of improvement justifies the increased cost and expense of additional testing.

b. Nasal aerosols

Lines 549-552: The last line of this paragraph (relating to reporting of individual spray data) should be deleted; reporting the mean of up to 3 consecutive sprays is sufficient.

3. Drug in small particles/droplets, or particle/droplet size distribution by cascade impactor

Lines 554-565: The requirement for cascade impaction testing has been clarified and is an improvement over the previous draft, but the requirement to use a multistage impactor is very prescriptive and hard to understand in the context of the data to be reported (sum of drug deposited on lower stages). If the basis of the test is really to address safety concerns regarding adverse pulmonary effects of excipients, would this not be better demonstrated by direct studies on these excipients at the kind of levels at which they might be inhaled, or greater specificity regarding the quality of input materials? We suggest revision of the opening paragraph of this section as follows: "The aerodynamic diameter of droplets or particles is an important factor in the deposition of drug in the nasal passages and should be measured using an inertial impaction method (e.g. multistage cascade impactor (CI)) or other validated technique. Analytical data should be based on a validated chemical assay. In the case of CI testing, we recommend that analytical runs include at least"

a. Nasal sprays: drug in small particles/droplets

Lines 569-574: We suggest revision of this text as follows: "Small droplets, for this test and dosage form defined as smaller than 6μ m, may potentially be delivered to regions of the airways beyond the nose. This test is intended to determine the amount of drug in small particles/droplets. If a multistage CI is used, then small droplets can be defined as smaller in size than the nominal effective cut-off diameter of the top stage. For example, for USP 25 Apparatus 1 (<601>), an eight stage...."

Line 581: We suggest inclusion of the following sentence: "The test for drug in small particles/Droplets for nasal sprays is not intended to provide PSD of drug or aerosolised droplets".

Line 584: We suggest inclusion of the following phrase "Measurable levels of drug in small particles/droplets would be a function of".

Line 587: The CI test can only be used for comparative purposes and cannot be used to assess a potential safety concern as the test system does not mimic the nasal cavity.

Lines 592-600: We suggest revision of this text as follows "Mass balance would be based on drug deposition on each of valve stem, actuator, adapters, induction port, any other accessories and the sizing apparatus and is recommended to be between 85 and 115 percent of label claim on a per actuation basis. The total mass of drug present as small particles/droplets is of primary interest. Therefore, in the case of multistage cascade impaction the pooled mass of drug deposited on all lower stages and filter can be reported".

b. Nasal aerosols: Particle/droplet size distribution

Lines 623-631: The guidance recommends that the determination of a profile be based on drug deposition at 11 sites and then goes on to state that the BE limit for the profile comparison depends on the number of stages and other accessory deposition sites. Does this imply that the sponsor would generate the profile at the 11 sites, but the sponsor has the option of combining stages/sites for performing the profile comparison? We suggest that a clarifying sentence be added that allows the sponsor to determine how to group stages and sites for performing the profile comparison or that the profile comparisons must be performed on the 11 individual stages. We assert that for BA/BE purposes, the comparison of T to R should be based on an assessment of individual stages and accessories, not groupings, as this can mask shifts in the PSD. Without the statistical appendices, it is impossible to evaluate the suitability of these measures for determination of BE.

4. Drug particle size distribution by microscopy

Line 639: The requirement for Drug PSD by microscopy in the spray following actuation appears to be a new requirement that does not appear in the recently finalized CMC guidance. If included in the final BA/ BE guidance this test should not be included as an additional routine release requirement.

Lines 647-650: The suggestion to evaluate PSD of drug substance in the presence of particulate matter from excipients for suspension formulations seems inappropriate as it may not be possible to differentiate between particle types, unless their shapes are very different. It is not clear what value this will add to the characterization if it is non-quantitative and other techniques have been used to assess the PSD of drug using a drug-specific method.

5. Spray pattern

Line 664-671: Terminology that is not in common use is introduced in this guidance. The agency should provide a glossary with specific definitions for Center of Mass and Center of Gravity, with photographs describing the interpretation of each term.

Lines 667-675: This guidance, or the companion CMC guidance, should confirm whether assessment of spray pattern by automated image analysis or manual methods will be acceptable for routine quality control testing of nasal spray products. We support the current flexibility included within the guidance to permit automated image analysis but do not believe this methodology should be prescribed.

Lines 704-708: Dialogue regarding suitability of nonspecific visualization reagents may be helpful.

Line 712: More detail should be provided to define the COM when the analysis is conducted manually. An example should be provided. It is not clear how COM can be used when determined based on an image and not drug mass.

Line 746: It would be helpful to clarify whether impaction systems or single actuations (or both) are preferred. It is our opinion that the quality of data obtained using impaction techniques based on the use of TLC plates is of significantly inferior quality to that which can be obtained using laser scattering techniques.

Lines 748-751: It is not clear why the guidance recommends determination of the spray pattern at 2 distances from the actuator orifice. Given the nasal geometry, the spray will have limited opportunity to form before impacting on the inner surfaces; therefore we query the relevance of testing spray pattern at both 3-4 and 6-7 cm from the orifice.

6. Plume geometry

Line 771: It is not clear how plume geometry measurements will establish *in vitro* BA. Clinically, the plume is never formed inside the nasal cavity. Tests performed at a short distance from the nasal actuator (SP and DSD) are better *in vitro* controls.

Line 776: It is not clear how one can quantitate a spray plume. Suggest the word "Quantitation" is changed to "Characterization".

Line 809-814: It is not clear what scientific data support the suggested analysis procedure. Reference to scientific articles should be provided to ensure the analysis method is broadly applicable.

Line 812: Is the guidance intended to specify that the recommended limits for BE for this parameter are 90-111%?

7. Priming and repriming

Lines 229, 459-467 and 816-850: There are several references to priming requirements, including recommendations for study designs, included in this guidance. We agree that priming and repriming data should always be required for an ANDA and that the Test and Reference should match in terms of number of primes, but we suggest that methodology is more appropriately addressed in the CMC guidance and that this guidance should concentrate on standards for *in vitro* BA and BE. We suggest that the information in lines 835-850 would have been more appropriately included in the CMC guidance since it relates to the specifications that will be set for the drug product. The relevance of this information to determination of BA/BE, and the standards to be attained, is not clear.

Line 843: It is not clear why the requirements for the priming comparison are tighter than the requirements for dose uniformity, especially considering the analysis is based on single actuations even for products where the label requires multiple doses. This requirement is a significant expectation escalation for comparison of doses.

VI. CLINICAL STUDIES FOR LOCAL DELIVERY

- A. General Information
- 1. NDAs
- 2. ANDAs
- **B. Clinical Study Batches**

C. Clinical BE Study Design and Subject Inclusion Criteria

Lines 909-910: It is unclear what is meant by the "lowest labeled adult dose". We do not agree that use of the lowest labeled dose in this study offers the best sensitivity. It is important to distinguish between statistically significant differences and clinically relevant differences. A study conducted with the lowest labeled dose that could not achieve a clinically relevant difference from placebo could demonstrate statistical equivalence between Test and Reference. We suggest that the study should be conducted at the usual adult starting dose. For example for Flonase® 200mcg QD is the starting dose with the option to dose 100mcg BID and maintain patients on 100mcg QD.

Line 910: 'Prime' products according to labeling instructions.

Lines 915-917: A perennial allergic rhinitis (PAR) clinical study should be the recommended indication for a test of equivalent efficacy. Although seasonal and perennial allergy patients experience the common symptoms of sneezing, itching, rhinorrhea and nasal congestion, studies have shown that the perennial sufferer, both allergic and nonallergic, experiences more severe and sustained nasal congestion (References 2, 3). This severe nasal congestion is attributed to the intensity and persistence of the late-phase inflammatory response promoted by chronic, unrelenting exposure to allergens such as house dust mite and animal dander (References 4, 5). As a result, perennial rhinitis is more difficult to treat than simple seasonal allergic rhinitis and is considered to be most effectively managed with an intranasal corticosteroid as the mainstay of treatment (Reference 6). A PAR study that is four weeks in duration also offers a longer treatment period by which to assess safety and tolerability compared with a SAR study which is only two weeks in duration. A PAR study also provides greater evidence of efficacy for perennial nonallergic rhinitis (PNAR) than a SAR study because of their similar symptomatology and pathophysiology.

Line 919: The lower the level of symptoms required at baseline the harder it is to show differences and easier to show equivalence between treatments for mean change from baseline. We recommend that a specific criterion is defined which ensures a moderate to severe patient population is selected (e.g. 50% of the maximum total nasal symptom score over the baseline period). We do not recommend using a placebo run-in however since it is especially difficult to establish a minimum level of symptomatology following

a placebo run-in period, as most patients will improve with placebo nasal spray over the run-in.

Line 925: Due to the high variability of the placebo response with nasal sprays a placebo responder would be difficult to define or identify in a study of an aqueous suspension. What will the placebo response be compared to if the placebo run-in is not preceded by a "no treatment" screening period where baseline symptom scores are collected? Should the response during the run-in period be compared only to the AM score collected at Visit 1 prior to receiving placebo? One option would be to require a 7-day run-in with no treatment to establish baseline, followed by a 7-day placebo, followed by a 14-day treatment. However this is not practicable for most studies in seasonal allergic rhinitis due to the short duration of allergen season. Absence of allergen would make it easier to show equivalence but would generate meaningless data. Additionally, to address this risk, a post-treatment run-out without use of test article or placebo would establish that symptoms are still possible. Another alternative would be to conduct this study in PAR.

We do not believe a long placebo run-in is appropriate. The vehicle placebo has been shown to produce variable degrees of improvement as large as 35%. It is important to establish that patients have a suitable degree of symptomatology prior to being assigned to treatment, and a placebo run-in will underestimate the severity of the patient's symptoms. Furthermore, since patients are improved as a result of the vehicle placebo run-in, there will be less opportunity for improvement once patients are assigned to treatment. This will inappropriately underestimate the degree of improvement of the active drugs compared to the placebo and increase the likelihood that significant differences between treatments will not be observed, i.e., that they are comparable. The intention of the responders analysis to address this may not be fully practical during the study of a seasonal condition.

Line 925-928. Without details of the acceptable equivalence range recommendation, it is not possible to determine how to power a study to show both significant difference to placebo and equivalence to active.

Line 935: It is not clear whether BE criteria need to be met for secondary (instantaneous scores) as well as primary (reflective scores). The entry criteria are based on primary endpoint only. We suggest that BE criteria should be met for instantaneous scores at the end of the dosing interval i.e., for a QD product, T and R should be equivalent on iTNSS 24-hour post dosing. We recommend that the document be circulated for review including the appendices, which should clearly state what BE criteria need to be met.

Line 938 (baseline definition): This document should be consistent with the guidance for "Allergic Rhinitis: Clinical Development Programs for Drug Products", relating to new products, which recommends to **also** analyze AM TNSS and PM TNSS separately. It is very odd that the day 15 AM assessment is not considered, as this is the last assessment for end-of-dosing interval. Fourteen days of treatment should result in 14 daily (24-hour) assessments that consist of 14 AM and 14 PM assessments.

The guideline recommends using 7 previous measurements to define baseline. We recommend that the same number of evening and morning scores be used to make up the baseline i.e. would also include evening of Day 4 in the calculation of the baseline.

Line 942: The guideline suggests that patients should be selected based on TNSS at both the screening visit and during the run-in period. We recommend that the TNSS inclusion criterion only be used during the run-in phase.

Line 949: The guideline recommends using 27 ratings to define on-treatment summary. Again we recommend using the same number of evening and morning scores to make up the on-treatment summary value e.g. also include morning of Day 15 in summary, as this is the last assessment for the end-of-dosing interval. Also AM rTNSS ratings are primarily assessing (previous) nighttime symptoms and PM rTNSS ratings are primarily assessing the daytime symptoms. Fourteen days of treatment should result in 14 daily (24-hour) assessments that consist of 14 AM and 14 PM assessments.

Lines 954-956: Unless the packaging components are identical there can be no appropriate blinding. It is critical that the products appear identical to patients.

Line 961: The guideline defines the per protocol population as "and had no protocol violations". Based on the next sentence, this should more strictly say "and had no protocol violations impacting assessment of BE/efficacy".

Line 963: We suggest that this reference to the 1988 guidance be updated to acknowledge more recent ICH guidances on the Common Technical Document and Statistical Principles for Clinical Trials.

Lines 958-970: It is not explicit in the guideline (although it is clear in the ANDA Checklist available from the FDA website, relevant text referenced below) that the required study sensitivity analyses for T and R are analyses versus placebo. It is also clear in the checklist that both of these analyses are required to be significant at the 5% level. This should be explicit in the guideline. In addition, there is no mention of the clinically relevant difference for the T and R comparisons relative to P (Placebo). If the study is sized based on the BE comparison of T and R then smaller than clinically relevant differences in TNSS could be detected vs. placebo at 5% level. It would therefore be reasonable to reduce the size of the placebo group relative to T and R groups. However, sizing of placebo groups needs to take into account the fact that sponsors need to show that both T vs. Placebo and Reference vs. Placebo are significant and this will impact power. According to the ANDA Checklist, the criteria for BE seem to be based on a 80-120% rule, however TNSS generally appears to be normally distributed on the absolute scale and does not require transformation. It would therefore be beneficial to define the limits of equivalence on an absolute scale. Alternatively, the document should describe methodology for converting analysis on absolute scale into a percentage change (e.g. use of Fieller's theorem). The guideline also states in Section VI, D that the analyses should be expressed in absolute terms rather than percentage change. This highlights a clear disconnect between the main text and the ANDA checklist. It is also not clear in the document why 80%-120% was chosen. Given that the study would involve clinical endpoint data it would be sensible to base the BE rule on clinical relevance. It also seems strange that these are not 80-125% as 80-120% are not symmetric on the multiplicative scale. We recommend changing the ANDA Checklist document to reflect criteria based on an absolute scale.

The populations used for equivalence (T vs. R) and superiority (T, R vs. P) comparisons are defined differently in the guideline with equivalence based on per-protocol and superiority based on intent-to-treat. If the aim of the superiority design were to validate the trial then we would recommend using the same population for both comparisons (i.e. per-protocol).

D. Clinical BE Study Endpoints

Lines 974-995: We agree in general with the study design for the Clinical BE study. However we have the following recommendations for improvement:

- Onset of action and time to maximal effect should be considered an endpoint for equivalency as this may be a point of differentiation between T and R. We suggest that there should be some check on the onset of action response curves during the study (14 days of treatment period), especially at earlier time points.
- PAR offers a more sensitive/consistent model than SAR
- It is difficult to fully critique the study design without the statistical appendices to provide sample size calculation and further define equivalency.
- Regarding safety assessments, detailed nasal examinations should be required for all studies.
- Overall evaluation of response to therapy by subject should be included as a secondary endpoint to support the symptoms scores.

Lines 982-995: We agree with the use of this symptom scale.

Lines 990-999: This doesn't specify how the study should be powered nor does it mention the use of confidence intervals for analysis. Without information regarding the statistical design of the study it is hard to provide constructive comments. Line 996 should specify 12-hour AM and PM total TNSS and give the number rather than say pooled, unless a consistent use of "pooled" is adopted.

VII. PK STUDIES FOR SYSTEMIC EXPOSURE

A. General Information

Line 1012: We agree in principle with the statements in this section but note that it is not clear what "sufficiently high" drug concentrations over an "adequate time" means. These terms should be defined. Also, valuable information about tmax and Cmax could still be obtained in a PK study even if concentrations at other times are not measurable. Furthermore, Cmax and partial AUC data at early post dose times may be more reflective of nasal deposition and absorption than later time points where plasma concentrations are more likely to reflect orally absorbed drug from the swallowed dose. This information should not be discounted.

It is acknowledged that dose-response relationships have typically been difficult to establish in a clinical efficacy setting for these products. A dose proportionality PK study to differentiate systemic exposure associated with different doses of Test and Reference products might offer a sensitive model for discriminating between products. We recommend that the guidance include clinical protocols for *in vivo* tests in humans which have proven sensitivity to discriminate between Test and Reference products for determining the bioavailability and bioequivalence of an intranasal dosage form which is intended to deliver the active moiety locally (21 CFR 320.24 (b)).

Lines 1015-1016: It should be specifically stated that sponsors must develop or find the most sensitive "state of the art" assay.

Line 1022: We strongly agree that a PK study would be preferable to a PD or clinical study for assessment of systemic availability of Test and Reference products.

B. Study Batches

Lines 1033-1036: We agree with this approach to selection of representative batches.

C. Study Design and Subject Inclusion Criteria

Line 1045: Again, it is extremely difficult to provide constructive comment in the absence of the statistical section and we recommend that these elements of the guidance be circulated for public comment before the guidance is applied to pending agency business or finalized.

Line 1053: We acknowledge the desirability of using plasma concentration-time profiles from BA and BE studies to evaluate systemic exposure for suspension drug products. In order to obtain measurable levels, it may be necessary to select multiples of the therapeutic doses. It would be helpful if FDA included advice regarding how many more sprays than the therapeutic dose are appropriate in order to "avoiding the possibility of alteration of the drug deposition pattern within the nose at higher volumes". It is true that both T and R would have their deposition patterns altered but it would not be biased in one direction. However, where oral absorption is the major source of systemic exposure the findings from such studies are unlikely to be informative about topical exposure in the nose. Such studies are already required for BA assessment of NCEs. We suggest that in the situation where the plasma levels are too low to measure at the labeled dose that the dose should be pushed to detectable levels to adequately characterize the PK profile. We believe that administration of several actuations, a study of dose proportionality and/or the conduct of a multiple dose study are preferable to conduct of a PD-only study and that the guidance should be clear that these options should be fully explored before considering default to a PD-only study.

Line 1061: Are there published data that address the significance of the time interval between doses and the subject head position during dosing? If so these should be referenced.

Lines 1069-1071: There is little guidance on what are considered to be "adequate" plasma levels.

D. Study Measures

VIII. PD OR CLINICAL STUDIES FOR SYSTEMIC ABSORPTION

A. General Information

Lines 1093-1097: Valuable information can be obtained from a PK study even if drug concentrations are only able to be measured over a portion of the dosing interval. For example, if Cmax and tmax for the products are appreciably different from one another, this is important even if concentrations at later times can not be measured. We strongly agree that a PK study should be preferred if at all possible.

Lines 1099-1100: We do not concur with the rationale for conducting a PD study instead of a PK study.

Clearly, for a drug for which therapeutic doses yielded systemic levels that were around or just below the limit of quantification, determination of comparable systemic exposure would not be meaningful evidence upon which to make a determination of bioequivalence. Again, this section references statistical analyses that are not included in the guidance, which undermines the value of public comment at this time.

Line 1105: At present this section only addresses study designs for investigating BE - no BE standards are provided. If a PD primary endpoint is appropriate, the study should be powered to reflect the inherent variability of that endpoint. For example, if urinary cortisol was selected as a PD endpoint, a clinically relevant difference should be prospectively determined and the study powered on that basis. Demonstration of a large cortisol effect using oral prednisone is not relevant.

Line 1112: The rationale for using an active control that will likely produce large decreases in cortisol far in excess of what would be anticipated following nasal administration should be justified. The clinical significance of smaller decreases in cortisol is unknown. It is more important to have a measurable PD effect with the Reference and Test products than to show an appreciable effect with an active control. Using a PD marker that is not sensitive enough to detect differences between products is similar to having an insensitive drug assay to measure PK. There may be a treatment difference that cannot be detected. It may be more appropriate to consider use of an inhaled corticosteroid as the positive control.

Line 1124-1125: It is not clear how to determine the number of subjects needed for the PD HPA-axis study. Will the study be an equivalence study between T+AC placebo and R+AC placebo? If so, what should be the equivalence range? What about T+AC placebo (or R +AC placebo) vs. P + AC placebo? It is also not clear what 'relative assessment of the HPA-axis' will comprise.

Lines 1125-1130: For BE to be established, pharmacodynamic endpoints must be clinically relevant and the length of the studies has to be sufficiently long to result in exogenous steroid exposure, should it exist.

We are not aware that these study designs have been validated as PD models of systemic exposure for patients with rhinitis, nor do we understand the rationale for conducting this study in allergic rhinitis patients purely as a compliance measure. It should be specifically stated that a crossover design is not suitable where the allergen load is not constant e.g. during a seasonal allergic rhinitis study. A crossover study will need to include extensive washout periods if onset of effect is a desired endpoint. Finally, crossover studies in people who are sensitized to allergen (season or challenge) in the first part of the study would be fraught with challenges due to the "priming" of the nasal mucosa and the potential protection from this by an active drug.

Line 1132: We agree that it is appropriate for sponsors to seek specific advice on the appropriate design for such studies once all efforts to conduct a PK study have been exhausted.

This guidance should include clear discussion of the relative standards for determination of comparable bioavailability between formulations, and the bioequivalence standards that would permit substitution of one product for another. The methods described do not include sufficiently sensitive standards upon which to reach a determination of bioequivalence.

Lines 1150 and 1180: Alternative active controls to prednisone (e.g. inhaled corticosteroids) should be considered to minimize the risk to subjects and maximize study sensitivity.

B. Clinical Study Batches

C. Clinical Study Designs and Subject Inclusion Criteria

Line 1155: The rationale for a 6-week PD treatment period should be explained. It is not clear what is required from the efficacy analysis for the clinical systemic absorption study - significance of each of T and R vs. placebo/just trends? We assume study size is to be based on an HPA axis endpoint.

Line 1167: We suggest clarification of this study design e.g. subjects will be randomized to one of four treatment groups:

T+ Active Control (AC) placebo R + AC placebo P + AC placebo P + AC

Active control placebo will only be administered during the same period as active control is administered.

Line 1173: We recommend removing the sentence, "The matching active control placebo would be dosed on days when the active control is not taken, including the run in period". The previous text states when the placebo active control is taken and we would not recommend administering placebo active control during the run-in phase of the trial.

Line 1174: Variability would likely be less if this study were conducted in human volunteers.

Line 1186: Serum cortisol is a more sensitive PD measure than urinary cortisol and should be the preferred measure. Measurements should be more frequent than every 4 hours during times when the rate of change is greatest.

Lines 1187-1189: The guidance states that 24 hour urine cortisol or plasma cortisol should be the primary endpoint for the BE study. It also recommends (but doesn't require) that patients be domiciled during the 24-hour assessment. It should be **required** that the patients be kept domicile during the 24-hour cortisol assessment. Based on the difficulties inherent in obtaining complete 24-hour urine collections on an **outpatient** basis the guidance should not be flexible on this point. Technically if 24-hour plasma were chosen as the primary endpoint the patients would have to be domiciled anyway since plasma collections would not be technically feasible on an outpatient basis.

D. BE Study Endpoints for Corticosteroids

Lines 1193-1195: The variability in urinary cortisols can be very great and we strongly recommend collection of 24 hour AUC plasma cortisol data in an in-patient unit.

Effects on urinary free cortisol or serum cortisols may not be a sufficiently sensitive measure of HPA axis function for the purpose of making a determination of bioequivalence or interchangeability, as evidenced by FDA's class labeling for inclusion in the Pediatric Use section of package inserts for intranasal corticosteroids i.e., "This effect [on growth] has been observed in the absence of laboratory evidence of HPA axis suppression, suggesting that growth velocity is a more sensitive indicator of systemic corticosteroid exposure in pediatric patients than some commonly used tests of HPA axis function". Clearly, once a determination of BE has been reached, the Test and Reference products will be used interchangeably in all patient groups, including pediatrics. The appropriate substitution of one product for another would be of particular importance in

children and adolescents when growth/laying down of bone is occurring and in the elderly where bone mass is decreasing.

We believe that conventional measures of HPA axis effects are not currently sensitive enough to detect potential systemic steroid effects in children by a T nor to determine equivalence to R and, therefore, we do not agree that a PD study in adults would support extrapolation of a determination of BE to this vulnerable patient population. If it proves impossible to conduct a sensitive, rigorous PK study in adults then a PD study in adults should be supplemented by a PD study in children using a sensitive model, such as growth velocity, to assure BE in this patient subgroup. Our recommendation is based on the findings of the special Advisory Committee meeting convened in July 1998, on orally inhaled and intranasal corticosteroids and growth in children because of growing concerns about potential effects of corticosteroids on growth. These concerns were brought forward by the results of a study of intranasal beclomethasone dipropionate in prepubescent children that assessed its effects on growth. This study demonstrated an unexpected and statistically significant decrease in growth velocity in the treated patients versus a control group, and, surprisingly, a failure of adrenal function testing to be predictive of this growth effect. The FDA concluded that conventional measures of HPA axis effects are not predictive of potential growth inhibition, as reflected in the current labeling for intranasal corticosteroids. As one of the outcomes of this meeting, the FDA issued in November 2001 a guidance for industry, "Evaluation of the Effects of Orally Inhaled and Intranasal Corticosteroids on Growth in Children" that outlines the suggested requirements for performing a long-term growth study. This represents a sensitive PD model for assessing the BA of these products in children.

IX. NUMBER OF RESERVE SAMPLES FOR BA AND BE TESTING

X. MULTIPLE STRENGTHS

A. Solution Formulation Nasal Sprays

Line 1266: Delete 'by Cascade Impactor'.

B. Suspension Formulation Nasal Sprays

Line 1282: It is not sufficiently clear that all *in vitro* tests should be performed for low (indeed all) strength suspension products. Consistent with the previous section of this guidance (X.A.) and to provide added clarity, the text 'or higher' should be added to this sentence.

Lines 1288-1300: Amend the text as follows: "BE conditions for the higher strength product would include comparative formulations and container closure systems, comparative *in vitro* data and comparative *in vivo* data. BE conditions for the lower strength product would include:

1. Documentation of BE for the high strength Test and Reference products.

2. Proportionally similar Spray Content Uniformity through Container Life between high- and low-dose test product and high- and low-dose reference product".

Line 1303: If an ANDA is submitted for two dose strengths at the same time, would it be acceptable to not do *in vivo* BE studies at low strength (if evidence of *in vivo* BE at high strength were submitted) or does this only apply for sequential ANDA's? This is not clear in this section of the guidance document and we recommend the document be updated to cover this scenario.

No assumptions can be made with respect to dose linearity or proportionality between different strengths of these suspensions, given the difficulties in differentiation of clinical effects (or systemic exposure) of different doses, as evidenced in the medical officer's review of Vancenase Double Strength, NDA 20-469 (Summary Basis of Approval, 1996, overview of efficacy, page 125: "..these studies were not able to differentiate a fourfold difference in dose exposure between active formulations...not considered adequate to determine equivalence of two formulations with the same total dose exposure").

XI. SMALLER CONTAINER SIZES

Line 1318: Any changes, including minor modifications to the container-closure system, should be reviewed with FDA.

GLOSSARY

Provision of a glossary of terms would be helpful.

TABLE 1

The use of the population bioequivalence approach has been shown to be insensitive to mean changes in formulation due to scaling to reference product variability. The FDA should carefully consider the choice of acceptance range if this approach is to be applied to *in vitro* testing and ensure that the choice of acceptance range for such assessment of equivalence protects public health.

APPENDICES

We do not agree that these are stand alone documents. This guidance should not be finalized until the proposed standards and endpoints to be included in this section have been released for public comment. Bioequivalency determinations should be based on the currently acceptable approach.

REFERENCES

- 1. Daley-Yates PT, Price AC, et al. Beclomethasone dipropionate: absolute bioavailability, pharmacokinetics and metabolism following intravenous, oral, intranasal and inhaled administration in man. Br. J. Clin. Pharmacol 2001; 51: 400-409.
- 2. Naclerio RM.The pathophysiology of allergic rhinitis: impact of therapeutic intervention. J Allergy Clin Immunol 1988; 82: 927-934.
- 3. Bousquet J, Gaudano EM, et al. A 12-week, placebo-controlled study of the efficacy and safety of ebastine, 10 and 20 mg once daily, in the treatment of perennial allergic rhinitis. Allergy 1999; 54: 562-568.
- 4. Naclerio RM. Pathophysiology of perennial allergic rhinitis. Allergy 1997; 52(suppl 36): 7-13.
- 5. Drouin M, Yang WH, et al. Once daily mometasone furoate aqueous nasal spray is as effective as twice daily beclomethasone dipropionate for treating perennial allergic rhinitis patients. Annals of Allergy, Asthma & Immunol 1996; 77: 153-160.
- 6. Mandl M, Nolop K, et al. Comparison of once daily mometasone furoate (Nasonex) and fluticasone propionate aqueous nasal sprays for the treatment of perennial rhinitis. Annals of Allergy, Asthma & Immunol 1997; 79: 370-378.