Equivalence Considerations for Orally Inhaled Products for Local Action—ISAM/IPAC-RS European Workshop Report

Carole Evans, Ph.D.,1 David Cipolla, M.S.,2 Tim Chesworth,3 Eva Agurell, Ph.D.,4 Richard Ahrens, Ph.D.,5 Dale Conner,6 Sanjeeva Dissanayake,7 Myrna Dolovich, P.Eng.,8 William Doub, Ph.D.,9 Anders Fuglsang,10 Alfredo García Arieta, Ph.D.,11 Michael Golden,12 Robert Hermann, F.C.P.,13 Günther Hochhaus, Ph.D.,14 Susan Holmes, M.S.,15 Paul Lafferty,16 Svetlana Lyapustina, Ph.D.,17 Parameswaran Nair, M.D., Ph.D.,18 Dennis O’Connor,18 David Parkins, Ph.D.,19 Ilse Peterson,17 Colin Reins, M.D.,20 Dennis Sandell, Ph.D.,21 Gur Jai Pal Singh, Ph.D.,22 Marjolein Weda, Ph.D.,23 and Patricia Watson19

Abstract

The purpose of this article is to document the discussions at the 2010 European Workshop on Equivalence Determinations for Orally Inhaled Drugs for Local Action, cohosted by the International Society for Aerosols in Medicine (ISAM) and the International Pharmaceutical Consortium on Regulation and Science (IPAC-RS). The article summarizes current regulatory approaches in Europe, the United States, and Canada, and presents points of consensus as well as ongoing debate in the four major areas: in vitro testing, pharmacokinetic and pharmacodynamic studies, and device similarity. Specific issues in need of further research and discussion are also identified.

Key words: inhaler, IVIVC, equivalence, goalposts, pharmacokinetics, pharmacodynamics, in vitro, device, patient handling

Executive Summary

THE EUROPEAN REGULATORY APPROACH TO DETERMINING EQUIVALENCE OF LOCALLY ACTING ORALLY INHALED PRODUCTS (OIPS) RECOMMENDS A STEPWISE DEMONSTRATION OF EQUIVALENT PRODUCT PERFORMANCE VIA IN VITRO TESTS, FAILING THAT—VIA PHARMACOKINETIC (PK) SAFETY PLUS PK PULMONARY DEPOSITION OR LUNG IMAGING STUDIES, AND FAILING THAT—VIA PHARMACODYNAMIC (PD) OR CLINICAL STUDIES. NORTH AMERICAN REGULATORY APPROACHES DO NOT ALLOW EQUIVALENCE DETERMINATIONS BASED

1 Catalent Pharma Solutions, Research Triangle Park, North Carolina.
2 Aradigm, Hayward, California.
3 AstraZeneca, Macclesfield, United Kingdom.
5 University of Iowa, Iowa City, Iowa.
6 U.S. FDA, Rockville, Maryland.
7 Mundipharma Research, Cambridge, United Kingdom.
8 McMaster University, Hamilton, Ontario, Canada.
9 U.S. FDA, St. Louis, Missouri.
10 Fuglsang Pharma, Rudolstadt, Germany.
11 Spanish Agency for Medicines and Health Care Products, Madrid, Spain.
12 Pearl Therapeutics, Raleigh, North Carolina.
13 Clinical Research Appliance (cr.appliance), Radolfzell, Germany.
14 University of Florida, Gainesville, Florida.
16 Medical Technology Consulting, Hemel Hempstead, United Kingdom.
17 Drinker Biddle & Reath, Washington, DC.
18 Boehringer Ingelheim, Ridgefield, Connecticut.
19 GlaxoSmithKline R&D, Ware, United Kingdom.
20 Pearl Therapeutics, Morris town, New Jersey.
22 AXAR Pharmaceuticals, Corona, California.
on *in vitro* data only. Specific requirements regarding *in vitro*, PK and PD studies also vary among Europe, the United States, and Canada.

From the technical perspective, there is a general consistency among regions in the scope of important *in vitro* tests (e.g., dose uniformity, aerodynamic particle size distribution, sameness of active pharmaceutical ingredients (APIs), similarity of excipients and devices), although specific acceptance criteria—where they have been explicitly stated—vary.

Special challenges are posed by multiple-API products (sometimes called combination products or fixed-dose combination products) and by consideration of pediatric populations.

**Introduction**

Over the past few years, several workshops have discussed approaches for demonstrating (bio)equivalence of OIPs that would be needed either for marketing authorization of a generic/second-entry product, or for justifying changes/line extensions for a brand-name/innovator product. In 2009, the Product Quality Research Institute (PQRI) sponsored a workshop “Demonstrating Bioequivalence (BE) of Locally Acting Orally Inhaled Drug Products.”(1) In 2010, PQRI in coordination with the Respiratory Drug Delivery Conference sponsored a workshop “Role of Pharmacokinetics in Establishing Bioequivalence for Orally Inhaled Drug Products,”(2) focusing particularly on the potential use of pharmacokinetics (PK) as the sole indicator of *in vivo* BE for locally acting OIPs. Later in 2010, a “European Workshop on Equivalence Considerations for Orally Inhaled Products for Local Action” was coorganized by the International Pharmaceutical Aerosol Consortium on Regulation and Science (IPAC-RS) and the International Society for Aerosols in Medicine (ISAM). The objective of this European Workshop was to continue the constructive dialogue among industry, regulators, and academic researchers started at the preceding two workshops but to focus it on the European views regarding OIP equivalence determinations. This article summarizes the discussions at this ISAM/IPAC-RS European Workshop (“the Workshop”).

In this article, the term Test product (T) refers to the OIP for which equivalence needs to be demonstrated, for example, a generic product or a modified innovator’s product. Analogously, a Reference product (R) refers to the OIP that serves as a basis for comparisons in equivalence determinations, that is, the original product.

This article reflects the current state of understanding and knowledge gaps in this field as revealed at the Workshop; it should not be viewed as regulatory advice, and it may not fully represent the positions of individual authors or of any organization with which they are affiliated.

**Overview of Regulatory Requirements for Demonstrating Equivalence of OIPs for Local Action**

The European regulatory perspective on demonstrating therapeutic equivalence of OIPs is described in a CHMP Guideline that was issued as final in early 2009 and came into force in August 2009.(3) This guideline applies to products that are used to treat asthma and chronic obstructive pulmonary disease (COPD), and is relevant for a variety of OIP dosage forms, including pressurized and nonpressurized metered dose inhalers (MDIs), dry-powder inhalers (DPIs), and solutions and suspensions for nebulization. In Europe, locally acting products such as OIPs do not meet the strict definition of a “generic medicinal product.”(4) Such products are commonly called “hybrid medicinal products” and their submission basis is described in article 10.3 of directive EC/2001/83.(5)

Overall, the EMA guideline sets general EU expectations for approval based on equivalence studies. Regardless of the legal submission basis discussed above, individual Member States have a sovereign right on a scientific basis to object to marketing approval of a product that they associate with lack of proof of sufficient risk/benefit balance. In the European Mutual Recognition or decentralized procedures, these concerns and differences could be resolved through an arbitration process leading to a common decision. In national procedures, however, EU Member States may have different final requirements. Furthermore, interchangeability/substitution (which has implications for reimbursement) is always decided at the national level. This gives each country the flexibility to consider such factors as their population’s needs, climate, historical data, etc.

As depicted in Figure 1, the European guideline describes a stepwise approach to demonstrating therapeutic equivalence using *in vitro* and *in vivo* data. In brief, the concept first uses *in vitro* comparisons, followed, if needed, by PK studies, and finally PD or clinical effect studies. If equivalence is demonstrated at any of these major steps, the subsequent studies are not required. Each step is discussed in more detail in the remainder of this article.

In the stepwise European approach, comparative *in vitro* data from a number of tests (e.g., such as from a multistage cascade impactor or multistage liquid impinger) may be sufficient evidence of equivalence, but only if the OIPs being compared meet all of the criteria in the European guideline (see later in this report for a complete list).

If comparative *in vitro* data fail to show equivalence,(3) PK studies comparing systemic exposure (safety) and lung deposition (efficacy) can be employed to demonstrate equivalence of OIPs. In other words, *in vitro* results failing to show

![FIG. 1.](image)
equivalence can be superseded by PK studies demonstrating equivalent exposure and lung deposition \textit{in vivo}. If the comparative lung deposition and PK parameters meet the acceptance criteria, then the PK results are acceptable to support drug approval despite the \textit{in vitro} differences between the T and R OIPs.

If \textit{in vitro} and PK testing do not indicate equivalence, then PD testing is required. At the Workshop, some participants pointed out that relative to PK and \textit{in vitro} data, PD data are more variable and potentially less sensitive to differences between products. They questioned why a sponsor would be asked to perform \textit{in vitro} and PK studies at all if the lack of equivalence in \textit{in vitro} and PK data can be superseded by the less discriminating PD or clinical studies. In response, European regulators explained that:

- PD studies are acceptable only if they are sensitive (i.e., can distinguish between two different doses—see later in this report for further discussion).
- If \textit{in vitro} and PK data do not demonstrate equivalence, the regulatory decision for OIPs may be based on PD/clinical studies for both pulmonary efficacy and systemic side effects, because differences in \textit{in vitro} and PK characteristics might not impact clinical safety and efficacy. Thus, equivalence of \textit{in vitro} and/or PK studies offer the opportunity to waive PD/clinical studies, but the lack of formal \textit{in vitro} and PK equivalence calls for PD/clinical studies.

\textit{In vitro} studies are the starting point for demonstrating equivalence in the EU. If \textit{in vitro} equivalence criteria are not met, PK studies would be performed. The reason for going through these steps in this order (and not initiating PD studies right away) is that \textit{in vitro} and PK results might show substantial product differences such that \textit{in vitro} studies would be unethical or unlikely to succeed. Significant differences in \textit{in vitro} and PK data may strongly argue for reformulation or redesign of the Test product. Conversely, continuation of clinical development might be justified and reasonable if \textit{in vitro} and PK studies do not indicate substantial product differences. Further, PK study outcomes provide guidance on the nature of PD and/or clinical study requirements (i.e., determine the need for additional safety or efficacy PD studies).

In contrast to the EU where a stepwise approach is acceptable, the U.S. FDA and Health Canada in practice require all aspects of comparative testing, including \textit{in vitro}, PK and PD (even though FDA has no formalized BE guidance for OIPs yet). For innovators working on product line extensions but currently allows use of PK data as a surrogate for PD efficacy; comparative PD efficacy studies are required in all cases. The rationale for the more conservative North American approach to pulmonary PK data appears to stem from concerns about the ability to quantify APIs in the systemic circulation resulting exclusively from pulmonary absorption and concerns regarding the insensitivity of PK studies to differences in regional drug deposition that could affect efficacy.

Health Canada issued a guidance for second entry short-acting beta agonists (SABA) in 1999, which does not accept “blood-level” studies “unless it can be shown that the analyte measured in the blood indicates what went through the lungs and its effect.” The Health Canada 2007 draft guidance for subsequent market entry of inhaled corticosteroids (ICS) for asthma recommends determining the systemic exposure via PK studies; it also provides for systemic exposure to be determined in a pharmacodynamic (PD) study by assessment of the effect on the hypothalamic–pituitary–adrenal axis (HPA) if the plasma levels are too low to enable reliable analytical measurement. No specific \textit{in vitro} acceptance criteria for bioequivalence purposes are included in the 2007 draft guidance, although that guidance requires “complete chemistry, manufacturing, and quality data” as well as “appropriate comparative data versus the Canadian reference product.”

The differences in regulatory requirements between the EU, Canada, and the United States are summarized in Table 1.

### Considerations for “In Vitro Only” Equivalence

The first step in the European guideline describes the possibility of \textit{in vitro} equivalence testing as the sole surrogate for establishing therapeutic equivalence of OIPs. The relative ease of operation, the high power to detect differences between products, and the relatively low analytical method variability make \textit{in vitro} testing an attractive option for establishing equivalence of OIPs compared to the other approaches (e.g., PK, imaging studies). Generic OIP approval based only on the \textit{in vitro} data is in principle allowed in Europe, but up to now, only a very limited number of such approvals have been granted. During the Workshop it became evident that solution pMDIs and nebulizer solutions may have the highest probability of success following this pathway. By contrast, the T and R DPIs rarely interface with the patient identically; for example, due to the mouthpiece shape or dispersion mechanism. Before a company considers an “\textit{in vitro-only}” submission in the EU, it should ensure that certain prerequisites are met, as summarized in Table 2, which also lists the hurdles that must be overcome in order to show that “pharmaceutical criteria for equivalence” are satisfied. A product must satisfy all of these \textit{in vitro} pharmaceutical criteria for equivalence; otherwise, \textit{in vitro} studies alone will not be considered sufficient for substantiating equivalence in Europe.
The comparative \textit{in vitro} studies must be performed according to a preestablished protocol, including proposed batch selection, cascade-impactor-stage pooling, and \textit{in vitro} acceptance limits. As these are frequently points of discussion during marketing authorization procedures in the EU, it is recommended that the protocol be preapproved by the appropriate regulatory agency.

Of all the \textit{in vitro} requirements, aerodynamic particle size distribution (APSD) comparisons attracted the most attention at the Workshop. The remainder of this section focuses therefore on the APSD test.

**APSD comparisons**

**APSD acceptance criteria.** Comparisons between two OIPs are commonly made using inertial impaction methods, specifically either the Andersen cascade impactor (ACI) or the Next Generation pharmaceutical Impactor (NGI). The EU guideline requires the comparison to be performed per impactor stage or justified groupings of stages. The proposed acceptance criteria are: the 90\% confidence interval of the T/R ratio must fall within 0.85–1.18 (where 1.18 is the inverse of 0.85) for all stages and/or groupings, although the applicant can propose and justify alternative limits.

The limits have been intentionally chosen to be restrictive, because of the limited availability of studies demonstrating a correlation between APSD data and \textit{in vitro} performance. These restrictive equivalence criteria may be challenging to meet when comparing individual stages between T and R, due to the high variability of APSD measurements, especially in the case of very low API amounts on stages. A simulation study presented during the Workshop\textsuperscript{(19)} demonstrated that a hypothetical product with between-batch and within-batch variability typical of approved OIPs could not be shown equivalent to itself when two sets of batches were compared following the methodology of the EMA guidance. This suggests that the acceptance criteria mentioned in the EMA guideline may be too stringent, at least for some stages/groupings. Taking the variability of the R as a starting point for establishing more realistic acceptance limits (and they may be different for different stages) was proposed as a better alternative. It could be instructional to see results of a study using real data for OIPs currently available on the market (to the best of the authors’ knowledge, such a study has not been conducted yet).

Another question related to the EMA APSD criteria is whether the acceptance limits should be applied to the T/R ratio of API depositions, or to the absolute differences in API deposited on each stage. Furthermore, it was argued that the best approach may be a metric that in a scientifically valid way combines the stagewise differences into a single parameter that measures goodness of fit between the T and R particle size distributions. An analysis of one such approach\textsuperscript{(20)} had been conducted within PQRI, with the conclusion that the chi-squared ratio metric is not sufficiently discriminating;\textsuperscript{(21–23)} however, no alternative recommendations were made by that

<table>
<thead>
<tr>
<th>Region</th>
<th>Regulatory basis for approval</th>
<th>Application type</th>
<th>Guideline/guidance</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU</td>
<td>“Stepwise” approach using comparative \textit{in vitro}, PK, and PD data; approval may be granted on evidence of equivalence at any step. PD safety and efficacy studies are the last step in cases where \textit{in vitro} and PK data are unable to demonstrate equivalence (e.g., PK-safety data shows noninferiority, instead of equivalence of T vs. R). National specifics may differ and may influence the interpretation of EMA guidelines.</td>
<td>Abridged application for “hybrid medicinal product,” also known as “hybrid application.” Variations/extensions to an approved application.</td>
<td>CHMP Guideline on the Requirements for Clinical Documentation for Orally Inhaled Products (OIP) including the Requirements for Demonstration of Therapeutic Equivalence Between Two Inhaled Products for use in the Treatment of Asthma and Chronic Obstructive Pulmonary Disease (COPD) in Adults and for Treatment of Asthma in Children and Adolescents, CPMP/EWP/4151/00 Rev.1\textsuperscript{(5)}</td>
</tr>
<tr>
<td>Canada</td>
<td>\textit{In vitro} comparative data are required but not sufficient for approval; PK safety studies, if feasible, are preferred over PD safety studies; otherwise PD safety studies are required. PD efficacy studies are required, with forced expired volume in the first second (FEV\textsubscript{1}) or other biomarker, for example, sputum eosinophil for inhaled corticosteroids (ICS). Use of PK data for demonstrating equivalent efficacy is currently a subject of debate in this field.</td>
<td>Generic drug application (ANDS); Bridging studies in New Drug Submissions (NDS); Supplement to New Drug Submission (SNDS).</td>
<td>Guidance to Establish Equivalence or Relative Potency of Safety and Efficacy of a Second Entry Short-acting Beta\textsubscript{2}-agonists Metered Dose Inhaler, February 1999.\textsuperscript{(7)} Draft Guidance Submission Requirements for Subsequent Market Entry Inhaled Corticosteroid Products for Use in the Treatment of Asthma, August 2007.\textsuperscript{(8)} Joint Health Canada and EMA Guidance for Industry Pharmaceutical Quality of Inhalation and Nasal Products, 2006.\textsuperscript{(9)}</td>
</tr>
<tr>
<td>USA</td>
<td>“Weight of evidence” approach; all aspects of comparative evidence are required, that is, Q1, Q2, device, \textit{in vitro}, PK and PD.</td>
<td>Abbreviated New Drug Application (ANDA). Supplemental New Drug Application (sNDA).</td>
<td>No formal BE guidance for OIPs but public presentations and articles—see text for references.</td>
</tr>
</tbody>
</table>
group due to the agreed constrictions on the scope and time for that project.

**Batch selection.** The unsuccessful result from the above-mentioned simulation study is not solely due to the difficulty of showing equivalence for stages with low load and therefore high variability. Another major reason is the between-batch variability, either due to batch-to-batch inherent differences or due to changes over time during storage. Because typically only three batches of T and R are studied in the in vitro equivalence experiments, there is a risk that all three batches of R come from one end of the normal range. The strategy for batch selection should be developed in consultation with the appropriate regulatory agency. This batch-selection problem is a consideration for all equivalence testing, not just APSD testing.

**Stage groupings.** If impactor stages are grouped for the data comparison, the EU guideline recommends at least four groups of stages, and expects the grouping to be justified by the expected deposition sites in the lung. In practice, grouping is typically done for adjacent stages and frequently
justified by linking the selected impactor groupings to sections of the airways (e.g., mouth/throat area, upper and lower parts of the lung). However, impactors differ in many aspects from the airways' and studies investigating in vitro–in vivo correlations are largely absent or contradictory and nonconclusive. The utility of these justifications is therefore debatable. Another aspect of discussion at the Workshop was the fact that groupings may hide information that may be relevant for the equivalence assessment. In practice, stages with very low amounts of drug substance are frequently grouped with other stages, because these low amounts are claimed to be irrelevant for the clinical outcome. Although this assumption seems to be reasonable, it is difficult to prove that a small amount of drug deposited on a stage is irrelevant. Another reason for grouping stages with low API load is that the relative variability typically is inversely proportional to the load, making the likelihood to demonstrate equivalence very low even when the T and R truly are equivalent.

During the Workshop’s breakout session on in vitro testing, participants discussed whether some forms of abbreviated impactor measurement (AIM) could be viewed as representing deposition in the throat, central, and peripheral areas of the lung—at least in one population group, for example, an average young, healthy Caucasian male. It was pointed out however, that although AIM in combination with efficient data analysis (EDA) may be more discriminating in detecting small differences in APSD (and thus useful for quality control purposes, to ensure batch-to-batch consistency of APSD), the in vivo relevance of such differences, as well as in vivo relevance of the selected AIM cut points, remain unproven and will need to be demonstrated for any given product.

### Alternative particle sizing techniques

The use of cascade impaction as a quality control tool has been well established since the late 1980s. However, an ACI, described in the European Pharmacopeia (EP), had originally been designed to collect airborne microorganisms and was only later coopted for characterization of pharmaceutical aerosols. Thus it was not developed with the end goal to maximize usability in the pharmaceutical laboratory, nor to possess in vivo relevance. Other impactors were developed specifically for particle sizing of pharmaceutical aerosols, including the next generation pharmaceutical impactor (NGI) and Marple-Miller impactor (MMI). Although both are specified in the USP, only the NGI is recommended by the EP. The NGI possesses a comparable number of stages to the ACI, whereas the MMI has fewer stages and so may not be as discriminatory in the very fine particle size range (although this conclusion is now partially challenged by the findings supporting the AIM/EDA concepts). Other acceptable EP sizing devices include Apparatus A (the two-stage glass impinger), Apparatus B (the metal impinger), and Apparatus C (the multistage liquid impinger), each of which has fewer “stages” than an ACI or NGI. In theory, more stages should provide more information about APSD. However, because the total amount of API is finite, fractionating the dose into more than four to five size ranges may degrade the quality of information.

Other sizing techniques, not involving impactors or impingers exist but have not been used for equivalence testing, although they might be useful for product development and characterization, especially for solution-based DPIs.

In vitro–in vivo correlations (IVIVCs) for APSD. Flows and throat's: in order to ensure consistent impactor performance in the size ranges of interest, the cascade impactor is operated over a limited range of fixed flow rates. Unfortunately, the stage cutoffs for a situation when an impactor is used with a realistic inspiratory profile have not been established. This means that standard impactors cannot be operated in a way that is reflective of the patients’ inspiratory profiles, which vary both in flow rate and time profile. This limitation may be of particular relevance for the comparison of DPIs, where differences in device airflow characteristics and dose release mechanisms may exist depending on the flow rate. In such instances, it may be more appropriate to use patient-relevant flow rates and flow acceleration to characterize inhaler performance. One way that this can be achieved is to use simulation equipment that mimics an average patient’s inspiratory profile through the inhaler, but also interfaces to the cascade impactor (CI) in such a way as to permit sizing at constant flow rate appropriate to the impactor. This approach has been adopted in such instruments as the Electronic Lung or the Pari simulator. In addition, modification of the impactor induction port to more closely model in vivo oropharyngeal deposition has also been investigated; such approaches include use of oropharyngeal casts or more idealized throats such as the Alberta Throat. It is possible to combine both realistic mouth–throat models and realistic inhalation flow rates and then size fractionate the aerosol exiting the throat in a cascade impactor at a constant flow rate. However, at present there is no standardized approach for such measurements, and many of these models may be too complex for either equivalence comparisons or routine QC testing.

Mathematical modeling: development of in vitro methodology to predict in vivo deposition remains of widespread interest. Although CIs are an accepted QC tool to characterize APSD, they were never designed to simulate deposition in the human respiratory tract, and in fact, the collection efficiencies for various ACI stages differ dramatically from those for regional lung deposition. APSD determination in a CI is predominantly based upon particle deposition due to inertial impaction whereas the situation in the human respiratory tract is much more complex and involves inertial, sedimentation and diffusion mechanisms. Once the aerosol is generated, the individual particles may also change in size or shape due to hygroscopic growth or evaporation during transit in the respiratory tract—but this is difficult to model or predict.

Several paths have been proposed to predict in vivo lung deposition from in vitro APSD data. An accurate determination of the in vitro APSD characteristics of the aerosol that reflects realistic in vivo conditions (taking into account flow rate, temperature, humidity, and airway geometry) is key. One can then, in principle, apply a mathematical model to predict in vivo deposition. At a minimum, the prediction should include at least three regions: the extrathoracic, bronchial, and peripheral or alveolar regions. Another scenario would be to use a “realistic and standardized” throat and lung (which do not exist at present), to quantify the
deposition in the various regions of the respiratory tract. Combinations of these strategies can also be envisioned. For example, as described earlier, a realistic inhalation flow profile and throat model could be used to quantify extrathoracic deposition and then a CI be used to measure the APSD. The amount of drug deposited in the CI gives a reasonable estimate of the total amount of API deposited in the lung but if prediction of deposition in various regions of the lung (e.g., the bronchia and alveoli) were desired, then a mathematical deposition model would be required in addition to CI data. Part of the mathematical consideration would have to be the different flow rates and other parameters affecting in vivo deposition. A further approach would be to correlate APSD data with in vivo PK outcomes ($C_{\text{max}}$, AUC, etc.) as a reflection of the total and regional lung deposition. Such studies have met with limited success.

Correlating APSD and imaging: total and regional deposition can also be measured using OIPs labeled with radio-tracers and imaging. However, these studies are limited in part by the resolution of the scanning equipment and the need to use a surrogate product (to create the radio-labeled product). In addition, small ventilatory changes during inhalation of the labeled drug aerosol, as well as changes in airway geometry due to disease, mean that the practical clinical relevance of such studies is limited.

Individual patient IVIVC? Matching in vitro APSD profiles does not ensure equivalent in vivo deposition, dissolution, and clinical effects. Even if an IVIV relationship were established for APSD, such a general correlation might not be valid for a given patient’s airways’ anatomy, physiology, and disease state. Predictions of in vivo deposition will only apply to an average healthy adult or child. The patient-device interaction can additionally cause significant variability in performance of OIPs and consequently “IVIVC” between individual patients.

Ultimately, what is important is the safety and efficacy of the deposited dose, and so the deposition information would need to be linked to information on the site of action and/or the location and distribution of the receptors of interest. Even if the in vitro and in vivo deposition profiles are shown to be the same between T and R, differences in the physicochemical properties between T and R drugs could result in changes in the dissolution profile in the lung, which has the potential to change its safety and/or efficacy. Thus, acceptance criteria describing allowable differences in APSD profiles may need to depend on the specific drug, disease, and disease state being treated.

Considerations for PK Studies in Establishing BE of Locally Acting OIPs

The objectives of PK studies for demonstrating BE of locally acting OIPs are twofold: (1) comparison of total systemic bioavailability (as an indicator of safety, to ensure that T and R do not differ in systemic side effects), and (2) comparison of the pulmonary bioavailability (as an indicator of efficacy). In the European stepwise regulatory approach, PK studies that demonstrate equivalent pulmonary deposition combined with adequate safety data may be considered sufficient to establish therapeutic equivalence. Note that the term “bioequivalence” is often used in EMA guidelines when discussing PK, and will therefore be used in this section.

The EMA requires double-blind, crossover PK studies with clinically relevant doses and strengths in the intended adult patient population. Studies in children are to be limited to the assessment of the systemic safety, whereas efficacy (pulmonary equivalence) is to be addressed by PD or clinical studies (see further discussion of this issue later in this report). Within PK studies, the EMA recommends standard BE methods and metrics ($C_{\text{max}}$, AUC). In addition, the time at which $C_{\text{max}}$ is observed ($t_{\text{max}}$) is included as an additional outcome parameter for OIPs, which is not generally mandatory for solid oral dosage forms in Europe except in those cases where the rate of absorption has significant clinical importance.

**Studies as a marker of systemic exposure (safety)**

There is a general consensus that PK studies assessing total systemic bioavailability are a valid surrogate for PD safety effects. This is in keeping with standard PK principles used to assess BE of T and R (e.g., for oral tablets). If T exhibits greater systemic bioavailability than R (i.e., T exposure is higher than R exposure). The European regulators explained, however, that PK studies may demonstrate that the observed PK differences lack clinical relevance. In the case of higher systemic exposure levels achieved with T, the EMA would therefore accept properly designed and sensitive safety PD clinical studies (e.g., 24-h plasma cortisol suppression), and may conclude equivalence if the observed PK differences between T and R do not translate into clinically meaningful differences in safety characteristics. It was suggested that widening the PK acceptance limits based on well-established PK/PD relationships could be an additional way to justify that the PK differences lack clinical relevance.

**PK studies as a marker of pulmonary deposition (efficacy)**

The European guideline requires demonstration of equivalent pulmonary deposition, as a surrogate for efficacy. This could be done either by lung imaging (gamma) scintigraphy of the radiolabeled OIP or by PK efficacy studies (Fig. 1). Note that lung imaging cannot replace PK safety (or PD safety, if needed) studies. Furthermore, neither PK efficacy nor lung imaging studies are allowed in children, so PD or clinical studies are needed to demonstrate equivalent efficacy in children, per the EMA guideline.

The PK studies for OIPs are quite different from the traditional PK bioequivalence paradigm. Although it is accepted in the scientific community that PK studies of OIPs are the method of choice to probe for differences in the systemic exposure, the ability of PK studies to identify differences in the pulmonary deposition characteristics (defined as the extent and pattern of pulmonary deposition of an
inhaled active substance) is still being debated. A key question is whether PK studies are suitable for comparing the efficacy of OIPs. This is because plasma drug levels may not necessarily represent the drug concentration at the target site within the pulmonary system and do not determine the effect at the end organ of interest, that is, the pulmonary effect. Rather, plasma concentrations of an OIP drug reflect end-organ deposition measured in a compartment downstream of the target organ.

In PK studies involving drugs that do not have appreciable oral bioavailability (BA) or that are performed under conditions that do not allow oral absorption (e.g., due to a charcoal block) of the swallowed fraction of an orally bioavailable drug, the lung is the only compartment from which the drug is absorbed into the systemic circulation. Under such conditions, systemic plasma concentration time profiles may provide information on the absorption processes in the lung (deposited fraction of the dose, particle size distribution pattern, pulmonary retention time, dissolution processes, and central-to-peripheral deposition pattern).

For OIPs with appreciable oral bioavailability, PK efficacy studies need to be designed in such a way that they allow assessment of the pulmonary deposition. Two methods were discussed at the Workshop: the charcoal block method, which prevents gastrointestinal (GI) absorption of the swallowed fraction of drug; and the “early” bioavailability method, which provides information about lung deposition if absorption from the lung is faster than from the GI tract.

The EMA allows pulmonary PK data to serve as a surrogate for efficacy, that is, demonstration of BE in a pulmonary PK study may preclude the need for subsequent PD efficacy studies. If pulmonary deposition is lower for T than for R, evaluation of comparative efficacy by PD or clinical methods is required. If pulmonary deposition is higher for T than for R, additional clinical safety data may be required, along with further comparative consideration of in vitro characteristics (such as comparative APSD, especially FPD data) in order to evaluate whether the observed PK differences can feasibly be associated with clinically relevant differences of regional lung deposition pattern.

Support for the European regulatory approach stems principally from the few published studies that simultaneously evaluated pulmonary PK profiles and PD efficacy of different formulations/doses of OIPs. These studies—focusing on β-2 agonists—demonstrated a relationship between increasing lung dose and greater pulmonary effects. Literature reviews incorporating cross-study comparisons have also concluded that a similar relationship exists. It is possible that a particular drug is at the plateau of a dose–response curve, so a greater or lower drug deposition in the lung may not necessarily mean changed clinical response; but this question will have to be resolved through PD or clinical studies, which are required by the EMA if equivalent pulmonary deposition is not demonstrated.

Several other publications suggested that for ICS as well, PK studies can reflect the local deposition of OIPs at the site of action. For example, PK studies of a beclomethasone dipropionate formulation showed higher pulmonary deposition in an HFA compared to the CFC product, in agreement with the protective effects found in PD studies. Likewise, similar protection against methacholine-induced airway hyperresponsiveness (methacholine PC_{20}) in asthmatic patients provided by HFA and CFC formulations of budesonide was mirrored by their equivalent PK profiles in healthy adult subjects.

Two other studies showed that PK studies are more sensitive than PD or clinical studies in identifying product differences. Daley-Yates reported that neither in vitro data nor a 12-week clinical study in asthma patients were able to demonstrate differences between two DPI products combining fluticasone propionate (FP) and salmeterol (SAL), whereas a PK study displayed differences. Similarly, the authors of a recent study comparing two DPI FP/SAL combination products concluded that their results “… underscore the difficulty in relying on pulmonary function to compare inhalation products intended to treat asthmatic patients,” as PK studies showed the higher discriminatory power.

Charcoal block. For the charcoal block approach, the requirement to validate the effectiveness of the employed charcoal block for a given API(s) was emphasized. Validation can be accomplished either by referencing published evidence, or by the applicant’s own (in vitro or in vivo) validation studies. At the Workshop, concerns were expressed that oropharyngeal mucosal absorption could confound PK-based pulmonary delivery assessments. This concern, however, was countered as being largely hypothetical considering the small oropharyngeal mucosal surface area (compared to the bronchial and GI tract surface areas) and the generally short drug residence time, given that oropharyngeal deposits are likely to be rapidly swallowed. Furthermore, mouth rinsing, which is recommended for all ICS-containing OIPs, would effectively minimize the fraction of the dose that feasibly can be absorbed via this route.

Another concern with charcoal block methods was the possibility that the ingested charcoal may not only prevent the absorption of the swallowed fraction of dose, but may lead, for certain drugs, to a more complex interference with their overall disposition, for example, by disrupting enterohepatic recycling of drug amounts undergoing biliary or intestinal secretion. It was argued, however, that in general, this mechanism should not compromise the comparability of pulmonary deposition, because it would likely affect the T and R formulations in a similar way.

Early bioavailability. In principle, the measure of early exposure by assessing partial AUC values (e.g., 0–30 min postdose) could be an alternative approach for estimating pulmonary drug exposure, and it has already been employed by the FDA in the approval of generic albuterol MDIs. However, it was stated at the Workshop that a simple and straightforward general translation of this approach to other compounds is not applicable, as different absorption properties (in particular, the rate of absorption characteristics) will require specific time windows to be established to accurately capture the fraction of the dose absorbed from the lung. Hence, it was generally agreed that the early-BF approach will need to be validated for each API and formulation in question, and will depend on their particular PK characteristics. It was further mentioned that for solid oral dosage forms, it had already been shown that early partial AUCs display higher variability as opposed to C_{max} values. This is a separate issue deserving careful exploration before any recommendations in this respect can be made.
Using PK parameters to describe pulmonary fate of OIPs

In general, data from PK efficacy studies can provide the following information about a drug’s pulmonary fate: (1) total pulmonary dose and absorption, (2) deposition pattern in the lung (e.g., central vs. regional), and (3) absorption rate and residence time in the lung (the latter parameter is not explicitly discussed in the EMA guideline).

There is no controversy regarding total pulmonary deposition, as numerous studies have demonstrated that PK is sensitive to differences in the deposited dose, and thus, PK is able to differentiate between two products that deliver different drug amounts to the lung.

PK data can reflect not only the total amount but also the rate of pulmonary absorption of drugs that are not orally bioavailable or where oral BA is precluded, for example, by a charcoal block. The rate with which the inhaled drug is absorbed from the lung into the systemic circulation may contribute to systemic side effects, but it may also provide information about APSD characteristics and the time of residence of deposited drug particles in the lung. These properties are known to affect the pulmonary efficacy and the degree of uptake that can be measured.

Using PK parameters to describe pulmonary fate of OIPs

In general, data from PK efficacy studies can provide the following information about a drug’s pulmonary fate: (1) total pulmonary dose and absorption, (2) deposition pattern in the lung (e.g., central vs. regional), and (3) absorption rate and residence time in the lung (the latter parameter is not explicitly discussed in the EMA guideline).

There is no controversy regarding total pulmonary deposition, as numerous studies have demonstrated that PK is sensitive to differences in the deposited dose, and thus, PK is able to differentiate between two products that deliver different drug amounts to the lung.

PK data can reflect not only the total amount but also the rate of pulmonary absorption of drugs that are not orally bioavailable or where oral BA is precluded, for example, by a charcoal block. The rate with which the inhaled drug is absorbed from the lung into the systemic circulation may contribute to systemic side effects, but it may also provide information about APSD characteristics and the time of residence of deposited drug particles in the lung. These properties are known to affect the pulmonary efficacy and the degree of uptake that can be measured.

Using PK parameters to describe pulmonary fate of OIPs

In general, data from PK efficacy studies can provide the following information about a drug’s pulmonary fate: (1) total pulmonary dose and absorption, (2) deposition pattern in the lung (e.g., central vs. regional), and (3) absorption rate and residence time in the lung (the latter parameter is not explicitly discussed in the EMA guideline).

There is no controversy regarding total pulmonary deposition, as numerous studies have demonstrated that PK is sensitive to differences in the deposited dose, and thus, PK is able to differentiate between two products that deliver different drug amounts to the lung.

PK data can reflect not only the total amount but also the rate of pulmonary absorption of drugs that are not orally bioavailable or where oral BA is precluded, for example, by a charcoal block. The rate with which the inhaled drug is absorbed from the lung into the systemic circulation may contribute to systemic side effects, but it may also provide information about APSD characteristics and the time of residence of deposited drug particles in the lung. These properties are known to affect the pulmonary efficacy and the degree of uptake that can be measured.

Using PK parameters to describe pulmonary fate of OIPs

In general, data from PK efficacy studies can provide the following information about a drug’s pulmonary fate: (1) total pulmonary dose and absorption, (2) deposition pattern in the lung (e.g., central vs. regional), and (3) absorption rate and residence time in the lung (the latter parameter is not explicitly discussed in the EMA guideline).

There is no controversy regarding total pulmonary deposition, as numerous studies have demonstrated that PK is sensitive to differences in the deposited dose, and thus, PK is able to differentiate between two products that deliver different drug amounts to the lung.

PK data can reflect not only the total amount but also the rate of pulmonary absorption of drugs that are not orally bioavailable or where oral BA is precluded, for example, by a charcoal block. The rate with which the inhaled drug is absorbed from the lung into the systemic circulation may contribute to systemic side effects, but it may also provide information about APSD characteristics and the time of residence of deposited drug particles in the lung. These properties are known to affect the pulmonary efficacy and the degree of uptake that can be measured.

Using PK parameters to describe pulmonary fate of OIPs

In general, data from PK efficacy studies can provide the following information about a drug’s pulmonary fate: (1) total pulmonary dose and absorption, (2) deposition pattern in the lung (e.g., central vs. regional), and (3) absorption rate and residence time in the lung (the latter parameter is not explicitly discussed in the EMA guideline).

There is no controversy regarding total pulmonary deposition, as numerous studies have demonstrated that PK is sensitive to differences in the deposited dose, and thus, PK is able to differentiate between two products that deliver different drug amounts to the lung.

PK data can reflect not only the total amount but also the rate of pulmonary absorption of drugs that are not orally bioavailable or where oral BA is precluded, for example, by a charcoal block. The rate with which the inhaled drug is absorbed from the lung into the systemic circulation may contribute to systemic side effects, but it may also provide information about APSD characteristics and the time of residence of deposited drug particles in the lung. These properties are known to affect the pulmonary efficacy and the degree of uptake that can be measured.

Using PK parameters to describe pulmonary fate of OIPs

In general, data from PK efficacy studies can provide the following information about a drug’s pulmonary fate: (1) total pulmonary dose and absorption, (2) deposition pattern in the lung (e.g., central vs. regional), and (3) absorption rate and residence time in the lung (the latter parameter is not explicitly discussed in the EMA guideline).

There is no controversy regarding total pulmonary deposition, as numerous studies have demonstrated that PK is sensitive to differences in the deposited dose, and thus, PK is able to differentiate between two products that deliver different drug amounts to the lung.

PK data can reflect not only the total amount but also the rate of pulmonary absorption of drugs that are not orally bioavailable or where oral BA is precluded, for example, by a charcoal block. The rate with which the inhaled drug is absorbed from the lung into the systemic circulation may contribute to systemic side effects, but it may also provide information about APSD characteristics and the time of residence of deposited drug particles in the lung. These properties are known to affect the pulmonary efficacy and the degree of uptake that can be measured.
salmeterol was emphasized, which often results in the situation that \(C_{\text{max}}\) is observed already in the first blood sample (frequently taken 5 min postdose because of the perceived need for multiple inhalations for the proper quantification of systemic LABA exposures). The potential implications of the regulatory acceptability of the “first-sample \(C_{\text{max}}\) data” in lung-deposition BE PK studies were discussed, but no consensus was identified. Attending regulators emphasized that efforts should be undertaken by the applicants to manage an earlier initial blood sample (e.g., at 2–3 min) to more accurately capture maximum exposure levels (i.e., \(C_{\text{max}}\)). However, the need for multiple inhalations to achieve quantifiable systemic exposure over a sufficient amount of time may still make it difficult to accurately capture \(C_{\text{max}}\) despite those efforts.

Information on the regional (central vs. more peripheral) lung deposition of OIPs can also be obtained from PK studies (specifically, \(C_{\text{max}}\) values), with appropriate modeling. For low pulmonary permeability and/or slowly dissolving drugs, combined with more central lung deposition characteristics, a more efficient ciliary removal of the centrally deposited fraction of dose can be expected. \(^{67}\) In turn, products delivering highly permeable and fast dissolving compounds and/or achieve more peripheral lung deposition to smaller airways are expected to result in higher \(C_{\text{max}}\) values\(^{68,69}\) due to faster absorption.

**Study populations**

**Patients versus healthy subjects.** The EMA OIP guideline suggests conducting PK studies “in the intended patient population.” This requirement was subject to debate at the Workshop because BE studies should be performed in a population displaying the highest probability for discrimination between \(T\) and \(R\). PK studies in asthmatic patients display higher variability than those in healthy subjects, because of disease-related changes/ variations in airway caliber and bronchial hyperresponsiveness. In line with these considerations, PK studies in asthmatic patients have been shown to display higher variability than those in healthy subjects. \(^{70}\) The actual degree of airway obstruction in asthmatic patients was shown to determine the total extent of pulmonary deposition as well as regional deposition pattern (i.e., predominant deposition in central airways at conditions of pronounced bronchial obstruction, indicated by low \(\text{FEV}_1\)). These observations are supported by another study in asthmatic subjects, which reported a significant positive correlation between \(\text{FEV}_1\) and pulmonary deposition of beclomethasone and mometasone furoate. \(^{70}\) The overall tendency toward a more central airway deposition in asthmatic patients, therefore, generally has a potential to compromise the proper characterization of product-related differences in their regional airway deposition characteristics (i.e., central vs. peripheral deposition).

Furthermore, studies in patients are often associated with ethical constraints (i.e., washout of established treatments) and numerous confounding factors. Overall, the available data suggest that PK studies in healthy subjects are likely to be at least as, or more, sensitive than patient studies to detect performance differences between different products. The use of healthy subjects avoids the potentially confounding influence of variable airways obstruction across study periods in asthmatic subjects. In summary, it was considered reasonable to preferentially conduct OIP PK studies in healthy subjects. Indeed, there is already regulatory precedent in Europe to allow studies in healthy subjects (see EMA’s homepage, information about past approved products, public assessment reports).

**PK studies in children.** Physiological and anatomical differences between airways of children and adults affect the handling and performance of pMDIs and DPIs. EMA therefore generally requires *in vivo* studies in children for the BE assessment. As challenges associated with the time coordination between pMDI drug release and the child’s inhalation are generally solved in children by using pMDIs in conjunction with spacers, pMDI approval in children is linked to coapproval of specific spacer device(s) that need to be included in *in vivo* studies in specific age groups (see below).

In contrast to pMDIs, DPIs have the advantage of automatically coordinated inhalation and drug release. However, performance of DPIs depend on device resistance and inspiratory flow rate characteristics, which again can result in product performance differences between pediatric and adult populations. For this reason, EMA requires a balanced enrollment and representation of children across the relevant age categories intended for use, for example, <2 years, 2–5 years, and 6–12 years.

The EMA specific recommendations for PK BE studies in children differ, however, from those in adults. \(^{71}\) PK studies are only recommended to evaluate the systemic safety in children, but not the pulmonary deposition, as is the case for adults. The EMA guideline states that pediatric PK efficacy studies are inappropriate because they might increase the experimental burden on the child while only indirectly implying efficacy. Some Workshop participants speculated that the EMA might also be concerned about methodological difficulties of charcoal studies with serial PK assessments in pediatric populations.

This conservative position regarding PK studies in children, based on ethical concerns regarding frequent and numerous blood draws in children, is understandable. These concerns could be overcome, however, for example, by less frequent sampling around \(C_{\text{max}}\), and/or by using population PK approaches, which use more children but fewer blood draws per child. \(^{72}\) The employment of noninvasive sampling approaches of other biological matrices (e.g., saliva) could be another solution. These alternatives were unfortunately not discussed in the 2009 EMA guideline.

The recommendation to employ PK BE approaches in children only for the assessment of safety but not efficacy, seems to contradict EMA’s position for other dosage forms, for which full PK studies in pediatric populations are frequently used, especially because PK studies can mitigate the need for long-term clinical studies.

Despite these reservations, participants discussed the necessity of pediatric PK studies. Test OIPs might differ in their device design and handling, as well as in the airflow dependence of their deposition characteristics. Considering that \(T\) and \(R\) might differ in their *in vitro* profiles, for example, APSDs at the clinically relevant flow rates, the incorporation of PK studies could become necessary for DPIs with flow rate-dependent deposition characteristics, in pediatric age groups that are unlikely to achieve the range of adult inspiratory flow rates.
Bioequivalence studies for products with two or more APIs are a challenge, because of the necessity of multiple comparisons and the difficulty of assessing the effects of individual active components in a product. During the Workshop, the participants discussed whether for multiple API products with comparable in vitro APSDs of both APIs, both components need to be compared via PK studies. Achieving formal PK equivalence for multiple API products under a variety of different conditions (with and without charcoal, with and without spacing devices for pMDIs, and at different flow rates for DPIs) is a significant challenge. It was noted that for products containing two active components, there would be in total at least eight PK parameters to meet BE acceptance criteria even without considering $C_{\text{max}}$ requirements (i.e., AUC and $C_{\text{max}}$ for each of the two APIs, with and without charcoal each). The chance of meeting the 0.8–1.25 acceptance criteria for each of the eight comparisons might be low. The Workshop participants highlighted these resource and statistical challenges, and encouraged regulators to consider them.

Some participants suggested that the systemic exposure of one active component could be reasonably interpreted as reflective/indicative for the pulmonary delivery of the second active component, as long as the reflective/indicative for the pulmonary delivery of the second active component could be reasonably interpreted as tors to consider them.

Bioequivalence studies for products with two or more APIs are a challenge, because of the necessity of multiple comparisons and the difficulty of assessing the effects of individual active components in a product. During the Workshop, the participants discussed whether for multiple API products with comparable in vitro APSDs of both APIs, both components need to be compared via PK studies. Achieving formal PK equivalence for multiple API products under a variety of different conditions (with and without charcoal, with and without spacing devices for pMDIs, and at different flow rates for DPIs) is a significant challenge. It was noted that for products containing two active components, there would be in total at least eight PK parameters to meet BE acceptance criteria even without considering $C_{\text{max}}$ requirements (i.e., AUC and $C_{\text{max}}$ for each of the two APIs, with and without charcoal each). The chance of meeting the 0.8–1.25 acceptance criteria for each of the eight comparisons might be low. The Workshop participants highlighted these resource and statistical challenges, and encouraged regulators to consider them.

Some participants suggested that the systemic exposure of one active component could be reasonably interpreted as reflective/indicative for the pulmonary delivery of the second active component, as long as the reflective/indicative for the pulmonary delivery of the second active component could be reasonably interpreted as tors to consider them.

PK studies with multiple API products

Bioequivalence studies for products with two or more APIs are a challenge, because of the necessity of multiple comparisons and the difficulty of assessing the effects of individual active components in a product. During the Workshop, the participants discussed whether for multiple API products with comparable in vitro APSDs of both APIs, both components need to be compared via PK studies. Achieving formal PK equivalence for multiple API products under a variety of different conditions (with and without charcoal, with and without spacing devices for pMDIs, and at different flow rates for DPIs) is a significant challenge. It was noted that for products containing two active components, there would be in total at least eight PK parameters to meet BE acceptance criteria even without considering $C_{\text{max}}$ requirements (i.e., AUC and $C_{\text{max}}$ for each of the two APIs, with and without charcoal each). The chance of meeting the 0.8–1.25 acceptance criteria for each of the eight comparisons might be low. The Workshop participants highlighted these resource and statistical challenges, and encouraged regulators to consider them.

Some participants suggested that the systemic exposure of one active component could be reasonably interpreted as reflective/indicative for the pulmonary delivery of the second active component, as long as the reflective/indicative for the pulmonary delivery of the second active component could be reasonably interpreted as tors to consider them.
A further difference in regulatory approaches between regions concerns the doses to be compared; for SABAs, Health Canada recommends the use of the “usually prescribed dose and doses higher than normally prescribed”—in safety studies and in efficacy studies, if required, whereas the EMA requires that the highest approved dose be evaluated in comparative PD safety studies. (3)

(Bio)Assay sensitivity

In attempting to rationalize the regional differences, it is useful to consider the fundamental requirements of a study intended to evaluate equivalence between two products. The principal and overriding requirement is that treatment effects measured in a study be sensitive to differences in delivered dose or potency, that is, the study must have assay sensitivity. (77) To assure that a study satisfies this basic requirement, it has long been appreciated that an equivalence study must include at least two doses of T and/or R OIPs (78,79) enabling a within-study assessment of dose–response. If there is no dose–response information, then the study may lack the sensitivity to differentiate formulations, and therefore no inferences regarding equivalence may be made.

Despite this fundamental tenet, the comparison of single dose levels of T and R alongside a placebo arm has been advocated in recent years as an adequate regulatory approach by Health Canada. (8,76) This approach provides assay sensitivity in the sense that it allows an assessment of whether the active treatments are effective. It does not, however, provide any information about whether a proportional active-placebo difference would be achieved with, for example, half or double the active-treatment dose and as such, appears to be suboptimal. The imposition of stringent acceptance criteria

---

Table 3. PD Study Requirements in the European Union, the United States, and Canada

<table>
<thead>
<tr>
<th>EMA</th>
<th>Health Canada</th>
<th>FDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform approach for all drug classes (3)</td>
<td>Class-specific approach.</td>
<td>Only SABA guidance issued (now withdrawn). De facto approach detailed below.</td>
</tr>
</tbody>
</table>

- **EMA**
  - All drug classes (3)
    - ≥ 2 dose levels of T & R in efficacy study.
    - Doses “on steep part of DR curve” in efficacy study.
    - 2 dose levels in safety study.
    - Highest approved and lower dose in safety study.

- **Health Canada**
  - SABA (7)
    - ≥ 2 dose levels of T & R in efficacy study.
    - “Supratherapeutic” dose (e.g., 4 puffs) if required (and lower doses) in efficacy study.
    - 2 dose levels of T & R in safety study.
    - Supratherapeutic and therapeutic dose in safety study

  - LABA (76)
    - Guidance under development.
    - 1 dose levels of T & R (& PBO) in efficacy study.
    - Lowest marked dose in efficacy study.

  - ICS (9)
    - 1 dose level of T & R (& PBO) in efficacy study.
    - Lowest marketed dose in efficacy study.
    - ≥ 1 dose of T & R in safety study.
    - Sponsor to justify dose selection in safety study.

  - Muscarinic antagonists.
    - No guidance.

- **FDA**
  - SABA (16)
    - ≥ 2 dose levels of T & R in efficacy study.
    - Doses in efficacy study comparable to HC approach (1, 2±3 puffs).
    - 3 dose levels in safety study (within cumulative dose design).
    - Supratherapeutic and therapeutic doses in safety study.
    - LABA/ICS/Muscarinic antagonists
    - No guidance.
for equivalence may reduce the risk of falsely accepting as equivalent a nonequivalent T product. However, it does not correct for the inability to make a rigorous within-study assessment of sensitivity to differences in delivered dose.

With respect to other considerations (discussed below) such as the doses employed, dose level separation, study duration, and the population enrolled (including the use of a population enriched in “responders”), all should be subervient to the fundamental requirement for assay sensitivity.

**Selection of doses to be studied**

Given the requirement for assay sensitivity, a suboptimal dose, for example, 50 µg FP (one puff of a Reference MDI, which has more than one puff per the approved minimal therapeutic dose) may be considered acceptable as one of the dose levels in a PD efficacy study (although the dose should not be a noneffect dose in the model employed). Similarly, a supratherapeutic dose level of salmeterol may be considered acceptable in a PD safety study.

With respect to dose level separation, the minimum separation required to distinguish doses is desirable. This will depend upon the model employed and the OIP under test, but doses separated by a factor of four may be required in many PD efficacy models, while a lesser dose multiple may suffice in classical PD safety models, at least for ICSs and β2-agonists.

Finally, because the lowest marketed dose of most Reference OIPs lies near the plateau of the dose–response curve, development of a half-strength T, that is, a dose level below the minimum therapeutic dose approved for R, may increase the ability of the study to differentiate doses, improve assay sensitivity and better define the dose–response curve. This can increase the precision of the dose–response characterization and, in turn, decrease the width of the confidence interval of relative potency estimated using the Finney bioassay or $E_{\text{max}}$ modeling allied to bootstrap statistical methods. The latter, which is not specifically recommended by EMA, is more suited for assessing dose–response curve and saturation of effect ($E_{\text{max}}$) following administration of active treatments (half-strength and full-strength) and placebo treatment. The development and use in clinical trials of half-strength presentations may be complicated, however, by several factors, such as unacceptable dose-to-dose variations in delivered dose when very small quantities of drugs are administered. Use of a lower strength dose therefore should probably be reserved for situations where marketed doses are so high on the dose–response curve that the PD equivalence study will be impractical without use of the lower dose strategy. There is no consensus on this approach at EMA, and a discussion with the appropriate regulatory agency is therefore recommended before proceeding.

**Clinical study model**

As for dose selection for an equivalence PD study, a clinical model that can differentiate formulations should be employed in preference to the one that is more relevant for clinical use but which is less discriminating between dose levels. In this respect, the development paradigm for a generic/alternate OIP is quite distinct from that for a novel inhaled drug.

**Efficacy pharmacodynamics.** Careful selection of an optimal outcome measure and the clinical study design in which it is measured, are critical to the success of an efficacy PD equivalence study. It is not sufficient just to identify significant differences between active doses and placebo. The selected clinical study model must be capable of identifying significant differences in responses to different active doses. Most often this is done using the Finney bioassay statistical methods (as recommended by the EMA guideline). The statistical power of an equivalence study using this method is a function of both the variability of the measured response ($s$) and the slope of the dose–response ($b$). These two factors do not function independently, but as a ratio ($s/b$). The smaller the value of this ratio, the more powerful the study. Thus, the question of how steep is “steep enough” can only be meaningfully answered in the context of variability of the response measurement in the study.

Two study models have been successfully used for the PD assessment of short-acting beta agonist equivalence (Table 4): careful assessment of bronchodilation caused by the beta agonist and β-agonist-induced inhibition of bronchoconstriction with methacholine or histamine. Although each of these models provides technical and logistical challenges, when done carefully, both can provide sufficient statistical power—that is, a dose–response that is “steep enough.” Assessment of long-acting β-agonists will likely be achievable with minor adjustments in these approaches, such as adjustment of the timing of response measurements to account for later occurrence of peak effects than is seen with short-acting β-agonists.

For inhaled corticosteroids, identification of a clinical study model that provides sufficient statistical power for a PD bioequivalence comparison has been much more problematic. Several outcome measures and associated study designs have been proposed, but it appears that very few of these studies have been used with success in PD efficacy equivalence studies.

**Safety pharmacodynamics.** Equivalent systemic exposure should be demonstrated through a PK safety study, if possible. If PK equivalence criteria are not met, then demonstration of comparable safety between T and R is required. The appropriate duration of a safety study should depend on the therapeutic class of the active substance, and should confer adequate sensitivity. However, there is no clear regulatory guidance regarding duration of assessment for single-dose and multidose studies, and dose levels (e.g., approved dose vs. supratherapeutic dose). The data should demonstrate noninferior safety of T versus R based on relevant changes in vital signs, cardiovascular assessments, biochemical parameters, and frequency of adverse events. (The EU guideline uses the term “no worse” in this context, although elsewhere in the guideline, the requirement is for “equivalent” safety). Table 5 summarizes design considerations for pharmacodynamic safety studies of classes of drugs formulated as OIPs.

**When are pediatric PD studies required, and why?**

There is a considerable ongoing debate as to pediatric requirements for generic OIPs, fueled in part by the lack of definitive data to guide appropriate recommendations. The EMA guideline differs between adults and children in terms of acceptability of PK efficacy data (and consequently, in the
requirements for PD efficacy studies). Specifically, in children, the guideline suggests that PK studies are not an acceptable surrogate to assess efficacy. However, there is no clear rationale to support this difference in approach to adults versus children. The guideline, therefore, advocates the provision of pediatric clinical data to support the approval of generic OIPs in pediatric patients in all but those (rare) cases where equivalence between T and R OIPs may be satisfactorily concluded on in vitro grounds alone.\(^{(3)}\)

The EMA guideline highlights several differences in anatomy and physiology between adults and children, which may in turn influence drug delivery, such as differences in airway caliber, breathing patterns, airway resistance, and inspiratory flow profiles. The guideline concludes therefore that extrapolating data from adults to children may, as a result, be difficult. However, although the differences noted in the guideline between children and adults are indeed valid, it is pertinent to question the relevance of these differences to the generic OIP context, that is, if two OIPs are shown to be equivalent in adults (in either a PK and/or PD model), what mechanisms could lead to their lack of equivalence in children?

### Table 4. Pharmacodynamic Clinical Study Models Establishing Equivalence

<table>
<thead>
<tr>
<th>Drug class</th>
<th>Clinical study model</th>
<th>Comments from Workshop</th>
</tr>
</thead>
</table>
| Short-Acting Beta-2 Agonists | BRONCHODILATION | • Advantage: Technically easier to do.  
• Disadvantage: Easy to fail if subjects are not well selected and study is not meticulously conducted. |
| | INHIBITION OF BRONCHOPROVOCATION (BRONCHOPROTECTION) | • Advantage: Greater probability of success.  
• Disadvantage: Technically more difficult to conduct.  
• Unanswered questions: Methacholine is the currently favored bronchoprovocation agent. Unknown whether study designs using AMP or mannitol could be more statistically powerful. |
| Consensus: Well-developed field, with defined pathways for successful studies. | | |
| Long-Acting Beta-2 Agonists | Same as for Short-Acting Beta-2 Agonists | • Need modification for time points of measurement (SABA peak effect at \(~30\) min; LABA peak may be later).  
• Need longer washout before study to avoid pharmacologic tolerance to LABA effect. |
| Inhaled Corticosteroids | IMPROVEMENT IN LUNG FUNCTION (in asthma suboptimally controlled at baseline) | • Advantage: Clearly clinically relevant.  
• Disadvantage: Concern about carryover between treatment arms prevents crossover design.  
• Advantage: Clearly clinically relevant; sufficiently steep dose–response.  
• Disadvantage: Technically difficult to conduct; large screen failure rate to identify suitable subjects. |
| Consensus: Less mature field where it is less clear which approach is optimal. | | |
| Debate: Which is the better model? Each has pluses and minuses. | ASTHMA STABILITY MODEL | • Advantage: Clearly clinically relevant; sufficiently steep dose–response.  
• Disadvantage: Technically difficult to conduct; large screen failure rate to identify suitable subjects. |
| | EXHALED NITRIC OXIDE | • Advantage: Easy to measure.  
• Disadvantage: Clinical relevance less clear than asthma stability model or classic study of improvement in lung function. Several confounding factors including smoking, intercurrent infections, nitrate content of food, etc. To ensure adequate dose-response, careful subject selection required. |
| | AMP, METHACHOLINE OR MANNITOL CHALLENGE | • Advantage: Easier to conduct than asthma stability.  
• Disadvantage: Clinical relevance uncertain. To ensure adequate dose-response, careful subject selection required. |
| | SPUTUM EOSINOPHILIA | • Advantage: Easier to conduct than asthma stability.  
• Disadvantage: In some cases, clinical relevance is questionable. Practical concerns include inability of some patients to generate sputum, and the requirement for manual processing of the sample, which makes standardization across centers challenging. To ensure adequate dose–response, careful subject selection required. |
Differences in external pMDI design that might impact patient–device interaction are relatively limited. Different pMDIs generally differ only modestly in terms of mouthpiece shape and diameter. Conversely, different pMDIs may interact with a given spacer in different ways; hence, the relative deposition from two devices may differ when compared with and without a spacer. The former comparison is obviously of particular relevance to pediatric use, and spacer data should always be provided to support a generic pMDI submission for a product indicated for use in children. If equivalence cannot be inferred from an in vitro comparison, the PK and possibly PD studies with a spacer will be required, per the EMA guideline. However, the Workshop participants questioned whether both tests should be required because one model seems sufficient. A discussion with regulators prior to implementation of a one-model approach is recommended.

Metered-dose inhalers in pediatric PD studies. Differences in external pMDI design that might impact patient–device interaction are relatively limited. Different pMDIs generally differ only modestly in terms of mouthpiece shape and diameter. Conversely, different pMDIs may interact with a given spacer in different ways; hence, the relative deposition from two devices may differ when compared with and without a spacer. The former comparison is obviously of particular relevance to pediatric use, and spacer data should always be provided to support a generic pMDI submission for a product indicated for use in children. If equivalence cannot be inferred from an in vitro comparison, the PK and possibly PD studies with a spacer will be required, per the EMA guideline. In addition, if the Reference product was not authorized in children, a full clinical development program is required for the Test if it is intended for pediatric populations.

It is debatable whether a spacer study relevant to children could be performed in adults. Differences in plume geometry, which could theoretically alter the relationship between two pMDIs in adults versus children in the absence of a spacer, are not relevant when a spacer is involved. Conversely, lower tidal volumes in children may theoretically increase aerosol losses (via impaction, adsorption, sedimentation, and coagulation) prior to emptying of the spacer, although even young children using a large volume spacer should be able to inhale the spacer volume within 10 sec or less. Overall, therefore, it is unclear whether the physiological differences between adults and children could alter the relationship between two pMDIs shown to be equivalent in adults both with and without a spacer. Several participants suggested that if PK equivalence was demonstrated for two pMDIs in adult studies both with and without a spacer, then clinical and PD pediatric studies were not needed. Unfortunately, definitive data are lacking, so no consensus recommendations could be made.
Dry-powder inhalers in pediatric PD studies. For DPIs, the situation appears more complex still. Available devices are far more varied than pMDIs in terms of external design and the patient–device interface.\(^{90}\) Internal resistance and flow rate dependency may also vary considerably between devices.\(^{91–94}\) With respect to the patient–device interface, it is plausible that the relative effects of two very different devices (e.g., Accuhaler and Turbuhaler) on oral geometry may differ; hence, drug deposition may differ depending on the population in which they are used, although definitive in vivo data in this respect are lacking.

Another scenario in which pediatric data are likely to be relevant, is where the relative in vitro performance of two DPIs varies across clinically relevant flow rate ranges. For example, a potentially important difference in performance between two devices may become apparent at lower flow rates applicable to young children but not at flow rates applicable to adults. An alternative to conducting studies to assess the impact of low flow rates on drug deposition/effect in pediatric subjects may be to use targeted low flow rates in adult subjects (controlled by using an inhalation profile recorder). Some Workshop participants suggested that with equivalent PK and in vitro data at low and high flow rates, further pediatric studies of DPIs should not be required (assuming no differences in the patient–device interface). However, there was no consensus on this issue. See also a discussion on pediatric PK studies in this article.

A further often cited consideration when discussing DPIs relates to the differences in operational handling between devices. Such differences may clearly be very relevant to the successful use of two different devices by children (and indeed by other patient subgroups). It is, however, debatable whether specific studies should be undertaken to compare drug deposition or PD effects when such devices are used in different populations under nonoptimal conditions, because the ability to optimally use a device should be an absolute prerequisite to a physician prescribing any device to a given patient.

In conclusion, although there is currently a lack of consensus on this issue and no relevant North American guidelines, there is a greater likelihood that additional PD data may be warranted in children during development of a generic DPI, although specific development requirements can only be validly assessed on a case-by-case basis. The greater the similarity between generic and reference DPIs, the less onerous requirements are likely to be.

### Demonstrating Device Design Similarity

**Device similarity and patient interface**

The design of an OIP device and the interaction between the device and the patient are integral to reliable and consistent in vitro drug delivery. When changes in device design occur, these factors should be considered for demonstrating equivalence to the original device. Determining what changes in the device design are of practical and clinical relevance, can be difficult due to the lack of general IVIVCs for OIPs. It is a challenge to discern, a priori, design features that could impact patient handling and clinical outcomes. Published equivalence studies for pMDIs and DPIs do not typically contain extensive device details. As a result, the degree of device design difference that is permissible between OIPs or two versions of the same product is an area of considerable debate. The discussion is further complicated by the diversity of inhaled device designs and the multitude of potential changes between T and R or the changes that can occur during a product life cycle for a given OIP.

The importance of considering device design similarity in the context of equivalence comparisons of OIPs is recognized by regulatory authorities in Europe, Canada, and the United States, although their approaches and pathways to regulatory approval for device changes are different.\(^{95}\) The EMA guideline states that the patient handling of T and R devices should be similar if equivalence is to be demonstrated by the in vitro route alone. In Canada, the regulations require that device design attributes be considered as they relate to functionality and potential impact on safety and efficacy. In the United States, where an aggregate “weight of evidence” approach to demonstrating BE has been recommended, it is recognized that similarity of device shape and equivalence in design and operating principles are expected to be important for patient use and therefore for the OIP overall equivalence.

Products showing equivalence under standardized in vitro test conditions or in a clinical testing environment, where device operation is closely monitored, may not necessarily demonstrate the same therapeutic outcome as in patients’ hands outside the clinic. Variations in sound, taste, or other physical sensations associated with the OIP administration may impact patient’s perceptions and, therefore, potentially their compliance with therapy. For example, upon the transition from CFC to HFA pMDIs, the “cold freon” sensation of CFC sprays disappeared due to the softer, warmer HFA sprays, and this was reported by the patients as “device malfunction.” Similarly, substitution of one DPI design for another can lead to patient mishandling if design features such as dose loading or mouthpiece geometry are modified from the original device.

### Approaches to demonstrating similarity of OIP devices

The 2010 IPAC-RS Device Change Survey,\(^{96}\) focusing on the design, material, and manufacturing changes, and the subsequent discussions at a breakout session of the 2010 ISAM/IPAC-RS Workshop\(^{97}\) provided an opportunity to gather information on current attitudes and practices among pharmaceutical and device manufacturing industries, regulatory authorities, and academia with regard to OIP device changes and criteria for device similarity. The Workshop participants discussed the rationale and type of testing needed to determine equivalence when well-defined device changes occur (e.g., material changes, device shape). The objectives of the Workshop deliberations were to (1) describe current views regarding device changes between T and R or postapproval; (2) facilitate a move toward consensus on appropriateness and type of in vitro and in vivo testing for device changes, especially in light of risk management, quality-by-design (QbD), and self-regulation applied to device changes; and (3) highlight areas where regulatory requirements may differ from what is perceived to be technically required.

Both the results of the Device Change Survey and discussions at the Workshop indicated that current approaches to implementing and justifying OIP device changes are diverse and somewhat arbitrary, and that testing was often
performed merely to meet perceived regulatory expectations and not technically warranted.

Significant themes that emerged included the need for regulatory interaction, incorporation of risk management, classification of OIP device changes, and better regulatory guidance or a best practice document prepared by the industry.

**Interaction with regulatory agencies.** Participants expressed a strong desire for further dialogue between industry and regulatory bodies to increase understanding of regulatory expectations related to OIP device changes. There is a striking lack of consistency in approaches to support device changes, as reflected in the IPAC-RS Device Change Survey results. Testing is often conducted not because of an assessment of the potential impact of a device change on the Critical Quality Attribute(s) of a product but because of perceived regulatory expectations. It was agreed that development of a rigorous, science-based approach may be needed to ensure that any critical impact(s) of a device design change could be assessed and addressed, and that increased self-regulation by manufacturers could make the development and approval of changed OIPs more efficient. Many participants also commented that it would be valuable to discuss specific challenges presented by OIPs with reviewers and regulatory agencies.

**Use of risk management.** Many Workshop participants had experience using risk management tools as a part of device development, and felt that these tools could play a greater role in the development of OIP devices specifically, although there was a lack of consensus on how they should be applied. Although these tools are based upon well-established medical device standards such as ISO 13485 & ISO 14971, participants were uncertain of how they could or should be used to manage OIP device changes; additionally, there was concern about the acceptability of these tools to regulatory agencies, because most OIPs are regulated as drugs, not medical devices.

The IPAC-RS Device Change Survey results and the Workshop discussion also suggested that there was an interest from industry in moving toward a QbD approach as the standard way of constructing regulatory submissions for OIPs. This approach is closely related to the Risk Management and Quality System approach that exists for Medical Devices in Europe and North America, and many contributors felt it would provide a platform to deal with future changes in a more scientific and pragmatic manner. Due to a lack of clarity about how this might work, however, there was some hesitancy to pursue a QbD path. Although many companies were using risk management tools in development of devices and assessment of changes, few were including such information in regulatory submissions (e.g., New Drug Applications, Marketing Authorization Applications), thereby using an opportunity to challenge the current paradigm. Industry experience with regulatory responses to QbD submissions for OIP devices is therefore limited.

**Classification of device changes, guidance, and best practice.** Developing a prescriptive guidance document addressing all types of device changes would be extremely difficult and of limited value due to the vast range of device technology that exists now and may emerge in the future. However, many Workshop participants felt that additional guidance focused on risk management and risk assessment of device changes would be of value. The ISO 20072 standard for aerosol devices was acknowledged as an example of a helpful guidance document that ensures that the critical aspects of OIPs are adequately considered, but places the majority of responsibility for the risk assessment and decision making with the manufacturer.

A guidance for device change management need not come from regulators. The participants suggested that industry could develop a document describing best practices, points to consider and consensus recommendations.

Some classification of device changes that would guide manufacturers’ efforts should, however, be either issued or endorsed/accepted by regulators to be of maximum value.

**Recommendations from Discussions on Device Similarity**

A variety of routes exist for achieving consensus on testing to determine bioequivalence of changed devices. Increased alignment within and between all stakeholders, including industry and suppliers, as well as regulatory and standard-setting bodies such as the International Organization for Standardization (ISO), the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), and the Global Harmonization Task Force (GHTF), would contribute to a more harmonized, consistent and transparent approach. Enhanced communication between and within regulatory agencies, such as between the FDA Center for Devices and Radiological Health (CDRH) and the FDA Center for Drug Evaluation and Research (CDER) could also facilitate further alignment.

Regulatory or “best industry practices” guidelines should be developed or refined to be (1) flexible rather than prescriptive; (2) based on risk management principles, such as per ISO 14971 and ICH Q9; (3) aiming to classify and consequently qualify a device change based on an assessment of its potential impact on critical quality attributes; (4) incorporating appropriate technical input; and (5) linking testing requirements to clinically relevant parameters.

These considerations should be taken up jointly by industry and regulatory bodies. Prospectively, these stakeholders should assess whether a classification system for ONDP device changes would help define the testing approach and regulatory pathway to support these changes and, if it would, how OIP device changes could be categorized, as well as who should decide when a scientific and technical justification for a change is adequate and acceptable. Additionally, stakeholders should consider what benefits exist to increased self-regulation of activities by industry, based on risk management and quality management systems (as practiced for non-OIP medical devices). Considerations for increasing industry self-regulation include: what would facilitate or hinder this process, what steps would be necessary for implementation of such a system, and what would be required for its recognition by regulators. Among potential benefits of increased self-regulation are improved and more consistent OIP performance, more robust quality systems, and better transparency.
**Table 6. Areas for Future Work and Discussion**

<table>
<thead>
<tr>
<th>Types of studies for demonstrating equivalence</th>
<th>Questions for further research and discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>&quot;In vitro only&quot;</strong></td>
<td>1. Should <em>in vitro</em> tests for equivalence comparisons take into account patient–device interface aspects? If so, how should such tests be designed, or how should patient–interface aspects be integrated into existing tests? How could discrimination be demonstrated? How should the acceptance criteria be determined?</td>
</tr>
<tr>
<td></td>
<td>2. Under what circumstances, and in what ways, are the differences in <em>in vitro</em> deposition patterns (APSD) predictive of differences in <em>in vivo</em> deposition?</td>
</tr>
<tr>
<td></td>
<td>3. Could APSD patterns be predictively correlated with <em>in vivo</em> PK and PD data? If so, how? What is that correlation (perhaps specific to the drug or drug class), and how robust is that correlation across patients?</td>
</tr>
<tr>
<td></td>
<td>4. Should the acceptance criteria for APSD comparisons depend on the specific drug and disease state being treated?</td>
</tr>
<tr>
<td></td>
<td>5. Should acceptance criteria for APSD comparisons vary depending on the fraction of dose deposited on a certain stage (e.g., are the limits of ±15% appropriate for a stage capturing only a negligible fraction of the total delivered dose (e.g., &lt;5%)?</td>
</tr>
<tr>
<td></td>
<td>6. Should there be specific <em>in vitro</em> tests for bulk API (e.g., to address micronization issues)?</td>
</tr>
<tr>
<td></td>
<td>7. What are suitable methods for testing similarity of APSD profiles (e.g., comparison stage by stage, entire profile, grouped stages, or other)?</td>
</tr>
<tr>
<td><strong>Pharmacokinetics</strong></td>
<td>8. Should there be additional specific <em>in vitro</em> criteria for excipients (because excipients may influence drug behavior <em>in vivo</em>)?</td>
</tr>
<tr>
<td></td>
<td>1. Should the PK acceptance limits be widened based on well-established PK/PD relationships?</td>
</tr>
<tr>
<td></td>
<td>2. Assuming that <em>in vitro–in vivo</em> correlations exist at least for the inhalable fraction, should the PK BE goalposts be scalable to account for the R batch-to-batch variability <em>in vitro</em> and/or <em>in vivo</em>?</td>
</tr>
<tr>
<td></td>
<td>3. Should PK studies using the early bioavailability method in principle be acceptable as alternative to charcoal-block studies, and what would be the validation requirements, for example, regarding definition of an optimized time window for accurately capturing pulmonary absorption?</td>
</tr>
<tr>
<td></td>
<td>4. For multiple API products: in case of a reasonable codeposition pattern of active ingredients <em>in vitro</em> (e.g., comparable fine particle fractions and similar stage-by-stage APSD), could it be acceptable to use only one active ingredient of the product as marker of BE in <em>in vivo</em> studies?</td>
</tr>
<tr>
<td></td>
<td>5. For multiple API products: Would it be acceptable to opt for one of the active components demonstration of PK equivalence, while opting for the other component a PD-equivalence approach?</td>
</tr>
<tr>
<td></td>
<td>6. What emphasis should be placed on the ability of PK studies to differentiate between regional lung deposition pattern, while commonly accepted PD or clinical efficacy end points (e.g., FEV&lt;sub&gt;1&lt;/sub&gt;) actually also do not provide this information?</td>
</tr>
<tr>
<td></td>
<td>7. What PK efficacy studies should be recommended for pediatric populations? For example, could sparser sampling around C&lt;sub&gt;max&lt;/sub&gt; population PK approaches, or other body fluids (e.g., saliva) be used to enable PK comparisons of pulmonary fate/efficacy of OIPs, thereby avoiding the need for long-term clinical/PD studies in children?</td>
</tr>
<tr>
<td><strong>Pharmacodynamics</strong></td>
<td>1. Should additional PD studies be required if PK studies demonstrate that the total systemic exposure (safety) is equivalent between T and R but pulmonary deposition (efficacy) is higher for T compared to R?</td>
</tr>
<tr>
<td></td>
<td>2. Should half-strength PD studies be required/accepted to demonstrate assay sensitivity in efficacy studies? What would be the quality requirements for such an investigational “half-strength” product?</td>
</tr>
<tr>
<td></td>
<td>3. Should supratherapeutic dose(s) be included in PD safety studies (i.e., with a dose/strength higher than the maximal approved for the Reference product)?</td>
</tr>
<tr>
<td></td>
<td>4. What should be the standard design for assessing ICSs for efficacy?</td>
</tr>
<tr>
<td></td>
<td>5. What should be the duration of PD safety studies for LABA, long-acting muscarinic antagonist (LAMA), and ICS OIPs?</td>
</tr>
<tr>
<td></td>
<td>6. What are the specific requirements for demonstrating PD equivalence for LAMA OIPs?</td>
</tr>
<tr>
<td></td>
<td>7. In pediatric PD safety studies of ICS OIPs, why is the parallel application of two PD-models of systemic safety required (i.e., both HPA axis and growth assessments) required, while in adults HPA axis assessment is sufficient?</td>
</tr>
<tr>
<td></td>
<td>8. Are spacer studies needed in children if an adult PD study with a spacer demonstrates equivalence between T and R?</td>
</tr>
</tbody>
</table>

(continued)
Areas for Further Work and Discussion

This report summarized the regulatory approaches, state of knowledge, and current understanding regarding equivalence determinations for locally acting orally inhaled products, as discussed at the 2010 ISAM/IPAC-RS European Workshop. The consensus positions were presented and discussed in the four major sections of this report (in vitro, PK, PD, and device similarity). Even though generally consistent recommendations exist in this field (albeit with some differences in the details of regulatory requirements among the European Union, Canada, and the United States), there remain areas with insufficient knowledge and therefore in need of further research and discussion. These open-question areas are listed in Table 6.

The ISAM/IPAC-RS Workshop participants expressed the hope that future research by interested stakeholders, and constructive discussions in public fora would address these open questions.

Acknowledgments

This article represents a report of the workshop’s discussions, and may not fully reflect the positions of individual authors or of any organization with which the authors are affiliated. The authors are grateful to all speakers and attendees of the ISAM/IPAC-RS Workshop, as well as to ISAM and IPAC-RS organizations for the support of this project. The authors also thank Julianne Berry (Merck), a member of the ISAM/IPAC-RS Workshop Organizing Committee, and ISAM/IPAC-RS Writing Committee, for her valuables contributions and helpful discussions.

Author Disclosure Statement

The authors declare that no conflicting financial interests exist.

References


38. Brindley A, Marriott RM, Sumby BS, and Smith IJ: The
37. Mitchell J, Bauer R, Lyapustina S, Tougas T, and Glaab V:
36. United States Pharmacopiea: General Chapter
35. Andersen AA: New sampler for the collection, sizing, and
34. European Pharmacopoeia 5.0: Chapter 2.9.18. Preparations
33. Byron PB, Cummings RH, Nichols SC, Poochikian G,
32. Tougas TP, Christopher D, Mitchell JP, Lyapustina S, Van Oort
30. Tougas TP, Christopher D, Mitchell JP, Strickland H, Wyka B,
29. Tougas TP, Christopher D, Mitchell JP, Strickland M, Bauer R,
16. United States Pharmacopiea: General Chapter <601>
15. Mitchell J, Bauer R, Lyapustina S, Tougas T, and Glaab V:
10. Mitchell J, Bauer R, Lyapustina S, Tougas T, and Glaab V:
7. Mitchell J, Bauer R, Lyapustina S, Tougas T, and Glaab V:
5. Mitchell J, Bauer R, Lyapustina S, Tougas T, and Glaab V:
1. Mitchell J, Bauer R, Lyapustina S, Tougas T, and Glaab V:


99. See www.fda.gov for more information on the FDA, CDER, and CDRH.

Address correspondence to:
Carole Evans
Catalent Pharma Solutions
P.O. Box 13341
Research Triangle Park
NC 27709
E-mail: carole.evans@catalent.com