EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

16 March 2024
May 2018 (Version 4.0)
EMA/CHMP/101453/2024
Committee for Medicinal Products for Human Use (CHMP)

Guideline on the requirements for demonstrating therapeutic equivalence between orally inhaled products (OIP) for asthma and chronic obstructive pulmonary disease (COPD)

Draft agreed by Methodology Working Party, Quality Working Party, Paediatric Committee/Formulation Working Party and Rheumatology and Immunology Working Party	December 2023
Adopted by CHMP for release for consultation	21 March 2024
Start of public consultation	12 April 2024
End of consultation (deadline for comments)	30 October 2024

This guideline replaces "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 use in the treatment of asthma in children and adolescents (CPMP/EWP/4151/00 Rev. 1)".

Comments should be provided using this EUSurvey form. For any technical issues, please contact the EUSurvey Support.

Keywords	*Guideline, inhalation, orally inhaled products, therapeutic equivalence, asthma,* chronic obstructive pulmonary disease *(COPD)*

Official address Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
Address for visits and deliveries Refer to www.ema.europa.eu/how-to-find-us
Send us a question Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000 An agency of the European Union

Guideline on the requirements for demonstrating therapeutic equivalence between orally inhaled products (OIP) for asthma and chronic obstructive pulmonary disease (COPD)

Table of contents

Executive summary

This guideline is the 2nd revision of the CHMP Guideline formerly called "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 use in the treatment of asthma in children and adolescents". It addresses the requirements for demonstration of therapeutic equivalence (TE) between orally inhaled products containing the same active moiety(ies).

It is now clarified that the demonstration of TE between OIP is based on a stepwise approach, where TE could be demonstrated *in vitro* if all *in vitro* requirements are fulfilled or else preferably by means of pharmacokinetics if equivalent systemic exposure (as a surrogate marker for safety) and equivalent lung absorption/deposition (as a surrogate marker for efficacy) is demonstrated in spite of some *in vitro* differences. It is generally not recommended to aim at demonstrating TE using pharmacodynamic or clinical endpoints as these are deemed insensitive. The text on how to apply pharmacodynamic and clinical endpoints is thus considerably shortened or deleted.

The section on children and adolescents is shortened and it is now said to be acceptable to apply the same age limits as for the reference product in many cases. The conditions for extrapolation of PK data from healthy volunteers to the full patient population are also described.

In the previous guideline there was also some general information on pharmaceutical forms which is now deleted.

1. Introduction (background)

Existing CHMP documents that discuss the clinical requirements for the development of inhaled products - Guideline on the clinical investigation of medicinal products for the treatment of asthma (CHMP/EWP/2922/01 Rev.1) and Guideline on clinical investigation of medicinal products in the treatment of chronic obstructive pulmonary disease (COPD) (EMA/CHMP/483572/2012 -corr1) - focus primarily on the clinical development of inhaled products containing new active substances. This guideline is directed particularly at the requirements for demonstrating TE between OIPs containing the same active moiety(ies) and used in the management and treatment of patients with asthma and/or COPD.

The guideline was first published as points to consider in 2004 and revised for the first time and became guideline in 2009. Since then, a number of Q&A documents have been published by Quality Working Party (QWP) and former Pharmacokinetic Working Party (PKWP). Over the years, practice has been formed with scientific advice and approvals of medicines based on documentation not fully in line with the guideline in force and there was thus a need to update the document reflecting current practice.

2. Scope

This document provides guidance on the requirements for demonstrating TE between OIPs, including both, single active substance products and combination products.

The guideline focuses on abridged applications, but the principles described may be applicable for any other applications that are based on demonstration of TE compared to a reference product, such as line extensions, variation submissions or during product development. Also, in the case that there is a need

to confirm similarity to a product for which literature data is available (e.g., well-established use applications), the same principles apply.

In vitro aspects relevant for the establishment of TE are described in this guideline, but reference is also given to the Guideline on Pharmaceutical Quality of Inhalation and Nasal Products (EMEA/CHMP/QWP/49313/2005). Both guidelines are written to complement each other and should always be read in conjunction.

3. Legal basis and relevant guidelines

This guideline should be read in conjunction with the introduction and general principles, part I and II of the Annex I to Directive 2001/83/EC as amended and other pertinent elements outlined in the EU and the International Council for Harmonisation (ICH) guidelines, especially those on:

- EMEA/CHMP/QWP/49313/2005 Corr: Guideline on the pharmaceutical quality of inhalation and nasal products (under revision);
- EMA/CHMP/QWP/BWP/259165/2019: Guideline on quality documentation for medicinal products when used with a medical device;
- CPMP/EWP/239/95: Note for guidance on the clinical requirements for locally applied, locally acting products containing known constituents.
- EMA/CHMP/158268/2017 Rev.2: Guideline on the clinical development of fixed combination medicinal products;
- EMA/CHMP/83033/2023: Questions and answers on data requirements when transitioning to low global warming potential (LGWP) propellants in oral pressurised metered dose inhalers.
- CPMP/ICH/363/96: Note for guidance on statistical principles for clinical trials;
- CPMP/EWP/QWP/1401/98 Rev.1/Corr**: Guideline on the investigation of bioequivalence;
- EMA/CHMP/138502/2017 Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development.
- CHMP/EWP/2922/01 Rev.1 Guideline on the clinical investigation of medicinal products for the treatment of asthma
- (EMA/CHMP/483572/2012 -corr1) Guideline on clinical investigation of medicinal products in the treatment of chronic obstructive pulmonary disease (COPD)

Clinical trials, including bioequivalence and pharmacokinetic (PK) studies, conducted in the EU/EEA have to be carried out in accordance with Directive 2001/20/EC. Trials conducted outside of the EU and intended for use in a Marketing Authorisation Application in the EU/EEA have to be conducted to the standards set out in Annex I of the community code, Directive 2001/83/EC as amended.

4. General considerations in the investigation of therapeutic equivalence

4.1. A stepwise approach

Therapeutic equivalence means that the efficacy and safety profile of the test and reference products is sufficiently comparable so that a clinically relevant difference between products can be reliably excluded. The demonstration of TE between OIP is based on a stepwise approach, where TE could be demonstrated *in vitro* if all *in vitro* requirements are fulfilled or else preferably by means of pharmacokinetics if equivalent systemic exposure (as a surrogate marker for safety) and equivalent lung absorption/deposition (as a surrogate marker for efficacy) is demonstrated in spite of some *in vitro* differences. It is generally not recommended to aim at demonstrating TE using pharmacodynamic or clinical endpoints as these are deemed insensitive.

The *in vitro* comparison between the test and reference products is described in section 5. The use of only comparative *in vitro* data is acceptable if the product satisfies all criteria as set out in section 5.1. Data on *in vitro* comparability should always be provided for assessment, also in the case that some criteria are not fulfilled.

PK studies aim at evaluating pulmonary deposition and total systemic exposure compared to the reference product. PK endpoints are considered valid surrogate markers to adequately predict similarity in the pattern and extent of deposition in the lungs and the systemic exposure and, thereby, equivalence in both efficacy and safety. PK studies should normally be conducted in healthy adult volunteers. To assess pulmonary deposition, absorption of the active substance(s) from the gastrointestinal (GI) tract, if significant, may be blocked with charcoal (absorption via lung only), whereas for total systemic exposure, absorption from both lung and GI tract must be taken into account.

To be able to demonstrate TE regarding efficacy between test product and reference product, the test product has to show equivalence in pulmonary deposition to the reference product for the active substance(s) as described in section 6 below. In order to demonstrate TE regarding safety it is sufficient to demonstrate that the systemic exposure is not higher than for the reference product.

4.2. Additional considerations

4.2.1. Spacers

Spacers are required to be available for use with all pressurised metered dose inhalers (pMDIs). They should always be considered when a pMDI is used by a child and might also facilitate administration for adults. Appropriate data to support the use of a specific named spacer with a pMDI containing a specific active substance or specific combination of active substances must be included in the dossier. Thus, for pMDIs, data presented to demonstrate TE, should be conducted with and without a named spacer. If available, a spacer recommended in the reference product SmPC should be used. If the spacer is to be replaced subsequently by an alternative spacer, appropriate data must be presented. Two studies need to be conducted with spacer. One study should be performed comparing the aerodynamic particle size distribution (APSD) at 30 L/min flow rate with a 2 second delay. The delivered dose over tidal breathing should be compared in a separate study using the most sensitive, relevant breathing pattern as described in Ph Eur 2.9.44. In the case that TE is demonstrated using *in vitro* data for either the comparison with or without spacer but not for both comparisons, it is only necessary to perform a PK study for the comparison which did not demonstrate TE using *in vitro* data.

In those cases where PK studies have to be conducted with and without spacer and with and without charcoal blockade, the study with spacer and with charcoal blockade could be waived if it is sufficiently justified that the spacer eliminates the fraction deposited in the throat.

4.2.2. Products for nebulisation

This guideline applies also for products for nebulisation although it is acknowledged that the performance of these is highly dependent on the nebuliser used. As for spacers, data should be presented for at least one named nebuliser. Nevertheless, when solutions or suspensions for nebulisation have the same qualitative and quantitative composition as the reference product, the comparison of the APSD can be waived if other physicochemical parameters, including the particle size and polymorphic form of the active substance of suspensions for nebulisation, are shown to be equivalent.

4.2.3. Suprabioavailability

In cases of local suprabioavailability, i.e., if the test product displays an extent of pulmonary deposition appreciably larger than the reference product, reformulation to a lower dosage strength may be considered, followed by PK studies demonstrating TE between the reformulated test product and the corresponding strength of the reference product. In this case, however, the potential risk of medication errors needs to be addressed as the metered or delivered dose as labelled would differ from that of the reference product. If necessary, additional measurement to minimize the risk should be provided.

4.2.4. Fixed combination products

For a fixed combination product of known active substances, TE should be demonstrated for each individual active substance. Assuming that one active substance meets the *in vitro* criteria for TE and the other active substance fails, both substances should be evaluated in the PK study(ies) and fulfil the criteria regarding TE. However, it would not be necessary to conduct an additional study with charcoal if the charcoal administration was only necessary for the substance for which *in vitro* equivalence had been demonstrated.

5. *In vitro* comparison

The characterisation of the *in vitro* properties is the first step in the evaluation and demonstration of TE between the test and reference products. All *in vitro* criteria, as specified in section 5.1, should be studied. If all these *in vitro* criteria are not fulfilled, progression to *in vivo* studies is needed. The *in vitro* characterisation and comparison are essential and should always be performed irrespective of whether *in vivo* studies are needed. Section 5.2 covers additional aspects that need to be addressed to support results from the *in vivo* study(ies).

5.1. In vitro criteria for demonstrating TE

The test and reference products should be compared in order to conclude on TE. The *in vitro* TE should be performed and evaluated based on a study protocol including methods of comparison and acceptance criteria. TE is sufficiently demonstrated if the applied test product fulfils all the following *in vitro* criteria compared with the reference product:

1. The product contains the same active substance (e.g., same salt, ester, hydrate or solvate).

2. The pharmaceutical dosage form is identical (e.g., pMDI, non-pressurised MDI, dry powder inhaler (DPI)).

3. If the active substance is in the solid state (powder, suspension): any differences in crystalline structure and/or polymorphic form should not influence the performance of the product (e.g., aerosol particle behaviour, *in vitro* dissolution with relevant conditions).

4. Any qualitative and/or quantitative difference in excipients must be adequately justified and deemed not to influence relevant Critical Quality Attributes and/or any aspect of product performance other than those that are covered by the comparison of the APSD (e.g. mouth/throat feel, taste, patients' compliance, or safety).

5. Handling of the inhalation devices for the test and reference products in order to release the required amount of the active substance should be similar.

6. For DPI and breath-actuated inhalers, the inhalation device should have the same resistance to airflow (within ±15%).

7. The target delivered dose should be similar (within ±15%).

8. The APSD should be similar.

Data from the complete APSD profile of individual stages of a validated multistage impactor/impinger method should be provided with a sufficiently sensitive analytical method. Comparison may be performed per impactor stage or with justified groupings of stages/particle sizes. Data from each separate impactor stage should always be presented even when the comparison is performed on stage grouping. For stage grouping the following requirements should all be met:

- The group of stages should be prespecified. The strategy may be set based on pilot *in vitro* studies.
- Grouping may only be made by merging nearby impactor stages based on fraction size and is only justified if needed to ensure that the substance content in each group is sufficient to allow accurate estimation of the amount. Therefore, grouping of stages is only acceptable for stages with low deposition (i.e., <5% of reference product delivered dose) to the nearby stage with lowest deposition as well as grouping of non-sized fractions.
- At least 4 non-overlapping groups of stages or particle size fractions with defined cut-offs and not more than 3 impactor stages in each group are expected to be needed in order to give a complete description of the APSD.
- The non-sized fractions (i.e., throat/induction port, pre-separator) and fine particle dose (FPD) should be evaluated and compared separately. The FPD should be divided over at least 2 groups of stages.

The APSD comparison should be presented as the 90% confidence interval (CI) for the observed ratio of the geometric means of test and reference product and similarity is concluded if the 90% CI is within the acceptance limit of ±15% (85.00-117.65%). In case of grouping, data on the corresponding individual stages should also be presented but a descriptive comparison is then sufficient. Other approaches of evaluation of similarity of the average APSD of the populations of test and reference products may be proposed based on the variability observed in the amounts deposited in the stages or group of stages within the reference product. These approaches should preferably be confirmed at preceding scientific advice.

For DPIs with a device that is influenced by patient inspiratory effort, the APSD comparison should be performed with three different flow rates (30, 60, and 90 L/min).

Acknowledging that the number of comparisons may be large, a comparison in one stage or group of stages not meeting the acceptance criteria might be acceptable as an exceptional case. Nevertheless, the number of batches and samples per batch investigated should be sufficient to minimise the risk for Type II-error. No systematic deviation by the active substance, the product strength, the flow rate or the particle size group is acceptable.

At least three consecutive batches of the test product and three batches of the reference product should be tested with a minimum of ten inhalers of each batch. If there is a high variability, a larger number of batches and/or more inhalers per batch needs to be tested. The batches of the reference product used in the *in vitro* equivalence comparison should be representative of the product on the market including consideration of different ages.

5.2. Additional in vitro data of relevance for in vivo studies

Unless all criteria in section 5.1 are fulfilled, *in vivo* studies are needed to demonstrate TE (see section 6).

The formulation used in the *in vivo* study(ies) needs to be described in detail. Differences in formulation, inhalation device and manufacturing processes between clinical batches and the drug product to be marketed should be justified and the criteria for comparative *in vitro* studies in section 5.1 above may be taken into consideration.

To support the *in vivo* studies the following pharmaceutical aspects are important considerations.

5.2.1. Flow rate dependency of dry powder inhalers

In those cases where TE of a DPI is intended to be demonstrated by means of PK studies in healthy volunteers, it is necessary to compare the flow rate dependency of test and reference product to decide if studies in healthy volunteers can be extrapolated to the whole patient population. Patients may have impaired inspiratory capacity as compared to healthy volunteers and thus differences in flow rate dependency may be a concern.

Unless otherwise justified, comparative in vitro data on flow rate dependency should be provided for DPIs at a minimum of four different flow rates over the range of 30 to 90 L/min. The flow rate dependency for the test and the reference product is considered similar if the evaluation of FPD demonstrate either no flow rate dependency or similar flow rate dependency.

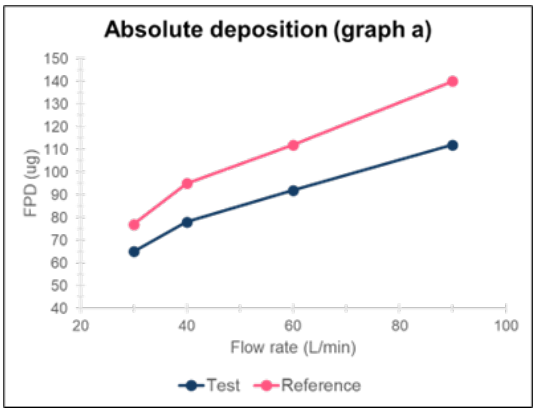
If there is a difference in flow rate dependency additional in vivo studies may be required (see section 6.3.2).

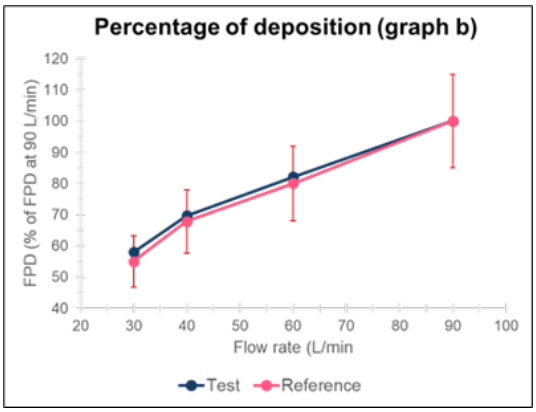
Test and reference products have similar resistance to airflow

If the resistance to airflow between test and reference devices differs not more than 15%, then the evaluation can be conducted using the flow rate. The following graphs are expected:

a. The FPD (y-axis) versus the flow rate (x-axis).

The percentage of deposition (FPD), where the FPD of the test and reference product at the flow rate of 90 L/min should be set as 100% (y-axis), versus the flow rate (x-axis).Example graph a:
Example graph b:





Similarity could be concluded if the point estimate of FPD of the test product in graph b is within ± 15% of the reference product for each tested flow rate (error bars in graph b).

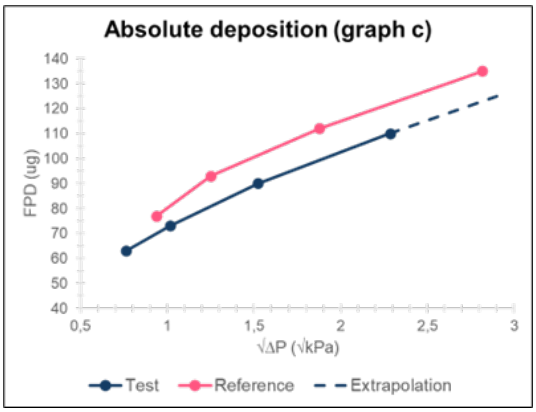
Test and reference products have different resistance to airflow

If the resistance to airflow between test and reference devices differs more than 15%, the evaluation should be conducted using the FPD (y-axis) versus the calculated √ΔP (x-axis) to allow for the comparison between test and reference product in a setting correctly mimicking the performance in different patient groups. The following graphs are expected:

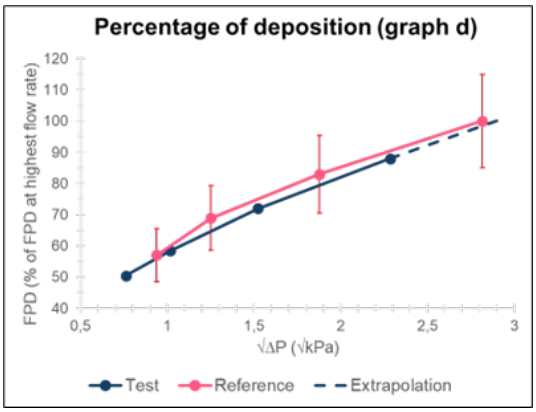
b. The FPD (y-axis) versus the square root of the pressure drop, √ΔP (x-axis).

c. The percentage of deposition (FPD) (y-axis) versus the square root of the pressure drop, √ΔP (x-axis). The FPD at the √ΔP corresponding to 90 L/min of the product with the highest resistance to airflow, should be set as 100% for both test and reference product. For the product with the lowest resistance to airflow the value of FPD set as 100% should be determined by extrapolation based on the slope of the graph between the last two points.

Example graph c:



Example graph d:



Similarity could be concluded if the interpolated FPD of the test product in graph d is within ± 15% of the reference product for each tested flow rate (error bars in graph d).

5.2.2. Investigation of several product strengths

In those cases where TE is demonstrated by means of *in vivo* studies with one of the strengths, *in vitro* proportionality should be investigated for both the test and the reference product across all proposed strengths to waive the *in vivo* demonstration with the additional strengths. To extrapolate *in vivo* data from one strength to other strengths comparable dose proportionality with test and reference product should be demonstrated.

If proportionality across all proposed product strengths is demonstrated with the test product, but not with the reference product, or vice versa, the two products cannot be deemed to be therapeutic equivalent for the strengths not studied *in vivo*. The test product must either be modified such that it matches the reference product or TE of the test product to the reference product should be established with more than one product strength and possibly with all product strengths, depending on which product strengths of the test product are not matched in respect of proportionality with the reference product.

In vitro proportionality should be demonstrated for the whole APSD although groups of stages could be used if a grouping strategy is justified (see section 5.1). The different strengths should be compared with a ±15% acceptance range in each stage. For products with a device that is influenced by patient inspiratory effort, e.g., DPI, the comparison should be performed at three different flow rates. If the different strengths of the test and the reference product are not shown to be proportional *in vitro* in the range of relevant flow rates, TE might be demonstrated by using a bracketing approach (see section 6.3.2).

5.2.3. Representative batches

Variability in APSD between batches of the reference product or within a single batch of a reference product through their storage period can be significant. Therefore, the batch(es) of the reference product used in the *in vivo* study(ies) should be representative of the commercial batches available on the market, including consideration for different ages or shelf-life of the product. The test product has to be representative of future batches and, therefore, the specification limits are critical to ensure similar characteristics even at the end of the shelf-life.

How the representative batch(es) is chosen should be fully discussed and justified. For some inhalers the APSD/FPD may change over time and in these cases ageing of the product should be considered. Characterisation of several batches of the reference product should be performed. A minimum of 5 batches may be sufficient if suitably justified. However, if the reference product shows great variability and/or degradation, a larger number of batches are needed. The FPD of the reference batch(es) chosen for the *in vivo* study(ies) should be as close as possible to the median of the observed values. A deviation within ±15% is reasonable.

6. Pharmacokinetics

6.1. Pharmacokinetic studies to investigate equivalence regarding safety (total systemic exposure)

In order to investigate systemic safety, the total systemic exposure for test and reference product should be compared in a PK study. The total systemic exposure is the sum of the absorption via the lungs and the intestinal absorption in a study where intestinal absorption is not prevented (i.e., in a study without activated charcoal blockade). Equivalent systemic safety can be concluded if test and reference products give rise to equivalent (or lower) systemic exposure (AUC_{0-t} and C_{max}), see section 6.3.3.

6.2. Pharmacokinetic studies to investigate equivalence regarding efficacy (lung deposition)

In cases where the contribution from the GI tract to the total systemic bioavailability following inhalation is negligible (<5%), or in the case that it is made negligible by active charcoal blockade, the area under the plasma concentration-time curve (AUC_{0-t}) is deemed a valid surrogate marker to reflect the amount of drug that has reached the lungs. As the rate of absorption from the inhaled particles is different at different areas of the lung, the deposition pattern within the lung affects the shape of the plasma concentration-time curve during the absorption phase, i.e., a relevant difference in deposition pattern can be assumed to be reflected in a difference in C_{max}. Thus, a difference in C_{max} between test and reference products may indicate that test and reference products are deposited in a different way in the lungs and absorbed at different absorption sites and thus that there is a difference between test and reference that may be clinically relevant.

The type of PK study that needs to be performed to investigate TE regarding efficacy depends on whether the contribution from the GI tract to the total systemic exposure following inhalation is negligible or significant.

6.2.1. Substances with negligible contribution from the gastrointestinal tract

For some orally inhaled medicinal products, the contribution from the GI tract to the total systemic exposure following inhalation is negligible (<5%) and a PK study without charcoal blockade can be used for both efficacy and safety comparisons. A low oral absolute bioavailability per se is, however, not synonymous with a negligible systemic contribution from GI absorption, since the contribution from the GI tract depends on the fraction of the dose being deposited in the lung and being swallowed, respectively, as well as on the fraction absorbed into the systemic circulation from each site. Reasons for the negligible contribution include poor intestinal absorption (e.g., chromoglycate, nedocromil), or an extensive first-pass metabolism (e.g., beclomethasone dipropionate, fluticasone, mometasone, ciclesonide).

6.2.2. Substances with significant contribution from the gastrointestinal tract

In this case there are two possible options as described below:

i. Study with activated charcoal

For drugs with significant oral bioavailability (e.g., budesonide, formoterol, salmeterol), a PK study with active charcoal can be performed to assess equivalence regarding efficacy. The charcoal blockade efficiency needs to be demonstrated (e.g., by using a method that has been shown to be effective in the literature).

ii. Early partial AUC in a study without activated charcoal

In the case that the absorption of the drug in the lung is very quick (e.g., median $t_{max} \leq 5$ min) and absorption occurs before the contribution of GI absorption is significant (e.g., salbutamol/albuterol, salmeterol, glycopyrronium, formoterol), $AUC_{0-30\ min}$ is acceptable as a surrogate for efficacy and AUC_{0-t} for safety. Thus, in this case, a study without active charcoal blockade is sufficient.

6.3. Design, conduct and evaluation of pharmacokinetic studies

6.3.1. General aspects

Pharmacokinetic studies intended to demonstrate TE between OIP should generally be performed according to standard methods for assessment of bioequivalence as described in the Guideline on the investigation of bioequivalence (CPMP/EWP/QWP/1401/98 Rev 1/Corr**). An open (bioanalytical laboratory blinded) study is acceptable.

6.3.2. Specific points to consider for OIPs

i. Study design

Generally, a single-dose cross-over study is recommended. It is critical that the sampling schedule is planned so that C_{max} can be reliably estimated and to avoid C_{max} being observed in the first sample post-dose. For example, formoterol and salmeterol have very rapid rate of absorption and thus early sampling is crucial in order to characterise C_{max}. Efforts should be made to have the first sample taken as early as possible (e.g., 2-3 minutes post-dose). It is however acknowledged that this is not always possible, especially if it is necessary to administer several inhalations due to low plasma concentrations and analytical limitations. The sampling schedule should also cover the plasma concentration - time curve long enough to provide a reliable estimate of the extent of exposure, which is achieved if $AUC_{(0-t)}$ covers at least 80% of $AUC_{(0-\infty)}$.

ii. Study population

Healthy adult volunteers generally demonstrate less variability in pharmacokinetics than patients. In addition, patients may be less discriminatory since lung depositions are mostly central in case of bronchoconstriction. Thus, the pivotal PK study(ies) should generally be performed in healthy volunteers.

For pMDIs (no flow rate dependency) and for DPIs in the case that the flow rate dependency for the test and the reference product is considered similar (see section 5.2.1), the study in healthy volunteers is sufficient.

If the flow rate dependency is not similar, TE cannot be concluded based on PK-data in healthy volunteers only but additional PK data showing equivalence at a low inspiratory flow rate (around 30 L/min) is needed. This study could be performed either in COPD patients with low inspiratory capacity or in healthy volunteers trained and monitored to inhale with low inspiratory effort or using an add-on device that increases the resistance to flow. Regular bioequivalence acceptance criteria should be applied. Unless equivalence can be demonstrated in a setting with low inspiratory flow rate, the extrapolation from healthy volunteers to patients of all categories cannot be confirmed and then no conclusion on TE may be drawn.

It is critical that all subjects included in a PK study are properly trained to inhale correctly in line with the product information and also to confirm during the study that subjects inhale correctly. If inhalation is not correctly performed, subjects should be excluded. Decision on exclusion should be made before bioanalysis.

iii. Choice of strength

If several strengths are applied for, it is sufficient to perform PK studies with only one strength, if dose proportionality *in vitro* is demonstrated for test and reference products (see section 5.2.2). If the

different strengths of the test and the reference product are not shown to be proportional *in vitro*, *in vivo* equivalence should be demonstrated with a bracketing approach. Bracketing should include the strengths most similar and most different from an *in vitro* perspective.

iv. Representative batches

The same batches should be used for the efficacy and safety PK study(ies), whenever feasible. Experience has shown that variability in aerodynamic particle-size distribution between batches of the reference product or within a single batch of a reference product through their storage period can be significant. There may even be situations where it may be difficult to demonstrate PK bioequivalence between batches of the same reference product especially in the case that a batch undergoes changes over time.

It is therefore critical that the batch(es) of the reference product used in clinical studies is representative of the commercial batches available on the market and that the test product is representative of future batches (see section 5.2.3).

In case of fixed combinations, it may be acceptable, if pre-specified in the protocol, to use different batches for each component to obtain representative batches for all active substances.

On very rare occasions, it may be difficult to find representative batches. The development of an IVIVC may be useful to correct the results of the PK study to justified parts of the APSD of the typical marketed batch of the reference product and the corresponding typical test product batch according to the proposed specifications (see section 6.4).

Another approach that might be acceptable is to show that the side batches (batches in the tails of the distribution) representing the test product specifications are not superior and not inferior to the side batches of the reference product obtained from the market.

6.3.3. Primary PK parameters to be analysed and acceptance criteria

The maximum concentration (C_{max}) and the area under the curve (AUC_{0-t}) should be evaluated. In the case that an early partial AUC ($AUC_{0-30\ min}$) is used as a surrogate for efficacy in a study without activated charcoal as described in section 6.2.2, this parameter is also primary and should be evaluated.

Therapeutic similarity with regard to efficacy can be concluded if the 90 % CI for the ratio of the test and reference product is contained within the acceptance interval of 80.00-125.00 for AUC_{0-t} and C_{max} (in a charcoal study or in a study without charcoal for a substance with negligible contribution from the GI tract) or for $AUC_{0-30\ min}$ and C_{max} (in a study without charcoal for a substance with very quick lung absorption for which an early partial AUC can be used).

To support safety, it is sufficient to demonstrate that the systemic exposure is not higher for the test product than for the reference product, i.e., the upper limit of the 90% CI for the ratio of the test and reference product for AUC_{0-t} and C_{max} should not exceed the upper bioequivalence acceptance limit 125.00%.

A widening of the acceptance criteria for C_{max} based on high intra-individual variability in line with the recommendations in the Guideline on the investigation of bioequivalence may be possible for substances where a wider difference in C_{max} is considered clinically irrelevant.

6.4. In vitro in vivo correlation (IVIVC)

As discussed in section 6.3.2 iv, the development of an IVIVC may be useful to correct the results of the PK study to justified parts of the APSD of the typical marketed batch of the reference product and the corresponding typical test product batch according to the proposed specifications in the rare occasions when it is difficult to find representative batches. Adjustment or normalisation may be acceptable if an IVIVC has been established previously between the *in vitro* parameters and the PK parameters for systemic safety and lung deposition and has been pre-defined in the study protocol. However, it should be noted that if a solid IVIVC was not established, normalisation will not be acceptable. The correlation should be shown for all actives in a fixed-dose combination product since the *in vivo* aerodynamic behaviour of the different drug particles may differ, although normalisation may be performed for one substance alone if the two products are considered similar for the other drug or no IVIVC is identified for that substance.

Due to inter-study differences, IVIVCs are expected to succeed only if they are investigated within a single study. It is essential to point out that different products at the same strength and dose with a different pattern of particle size distribution (PSD) should be included in the IVIVC.

The Applicant should justify the approach employed to establish an IVIVC, the selected method of normalisation and the criterion to define specifications based on the IVIVC. For example, the normalisation could be performed transforming the PK data to results expected for a "representative batch".

To support the conclusion of comparable pharmacokinetics, test and reference products may require independent normalisation according to their individual IVIVC relationships (as they are likely to be different from one another).

7. Pharmacodynamic and clinical studies

Endpoints as described in this guideline are deemed the most sensitive to detect differences between test and reference products and thereby the most relevant to use when demonstrating TE. In the case that data do not fulfil the acceptance criteria for PK endpoints, it is generally recommended to reformulate the product. Only exceptionally TE will be deemed possible to be established without being demonstrated kinetically, e.g., it could be applicable for some β_2-agonists.

If, however, other approaches with pharmacodynamic or clinical endpoints are considered, the study designs must be such that assay sensitivity is clearly shown at an acceptable level. It is acknowledged that for some active substances, and fixed combinations of such, appropriate study designs do not exist, but a full clinical data package would need to be provided instead of taking the TE approach.

Appropriate endpoints for TE efficacy are measures of airway function and/or inflammation, and appropriate endpoints for safety are measures of relevant biochemical and/or physiological parameters. Safety assessments including monitoring of adverse events should always be included in the efficacy studies regardless of design.

Regardless of the aim of the study, it is necessary to demonstrate that the sensitive part of the dose-response curve for the PD parameter under investigation has been studied. To allow for estimation of assay sensitivity, it is essential to include at least one non-zero dose level besides the level primary investigated.

As for the PK studies (see section 6.3.2), the same batch of reference product should be used for safety and efficacy PD studies, unless adequately justified, and should be representative for the

product on the market (see section 5.2.3). When feasible, it is of value to have access to PK data from the PD studies.

To conclude on TE in studies with PD or clinical endpoints, it is recommended that statistics is applied allowing for calculation of relative potency. The relative potency of the test product to the reference product is defined as the dose of the test product that produces the same biological response as one unit of the dose of the reference product. This analysis should be conducted based on the approach by Finney (1964)[1] for the primary efficacy variable, unless otherwise justified. The acceptance criteria for the 90% CI of the relative potency should be prespecified and normally retained within 0.67 to 1.50. This is as to support TE it must be clearly shown that a certain strength of the test product is more similar to the same strength of the reference product than the closest adjacent differing higher or lower strength (anticipated to differ by a factor 2 irrespective of whether there is an approved such strength or not). Any other choice of statistical approach must be sensitive enough to ensure assay sensitivity at this level.

8. Children and adolescents

In case of a new inhalation device, previously not approved for children, data on usability needs to be provided (see section 9). The characteristics of the delivery device may be such that the device is more difficult for a child to use than it is for an adult and, therefore, the child is less able to use the device correctly, or the child may use the device differently from an adult. Such differences in the handling of the product by a child may result in a changed risk/benefit relationship in the child compared with that seen in the adult.

In the case that it has been shown that the device can be correctly handled and emptied by children and the *in vitro* criteria for TE have all been fulfilled (see section 5.1, above) the age limit for the test product could be set at the same as the reference product without further data or justification. In case of pMDIs, the comparison should be made with the same spacer for test and reference products.

PK data generated in adults is deemed applicable supporting TE for adolescents (>12 years of age) without further justification. If the reference product has a lower age limit than 12 years of age the applicant is expected to provide a justification that the results of the PK study in adults can be extrapolated to the paediatric population. A prerequisite for extrapolation of PK data from adults is nevertheless that similar flow rate dependency has been demonstrated (see Section 5.2.1.) or that an additional PK study has been provided investigating exposure at a low inspiratory flow (see section 6.3.2.)

9. Usability studies

For medicinal products where the medical device and/or device part and the medicinal product form an integral product that is not reusable (hereafter called integral), a formal usability study (also named human factor study) may be required to demonstrate safe and effective use of the integral medicinal product by the intended user population as stated in the 'Guideline on quality documentation for medicinal products when used with a medical device' (EMA/CHMP/QWP/BWP/259165/2019), section 5.4.

Study participants should be recruited to include a number of distinct user groups including asthma and COPD patients (adults, and where appropriate children and adolescents) and caregivers, within

[1] Finney DJ. Statistical methods in biological assay. London: 104:1057–61. Griffin, 1964

which both reference product-naïve and experienced users should be included. A minimum of 15 participants should be recruited in each distinct user group.

Participant recruitment for these studies should aim to be representative of the intended user population incorporating general population trends (e.g., left handedness, elderly, patient with manual coordination difficulties, e.g., arthritic patients).

The study protocol should direct participants to simulate the use of the new device to deliver doses as per normal use (inhalers should be empty and participants should not be asked to inhale). The exercise should include the unpacking of a new inhaler from the patient pack, simulated delivery of the first dose, through the intended storage of the inhaler. Participants should be asked to simulate the delivery of further doses in order to assess the user interface with the inhaler through its life. Areas of focus should include ensuring the user understands key features of the device.

Clear acceptance criteria should be detailed together with rationale in the pre-specified protocol.

The outcome of this summative usability study should be reported through a usability report, which should include details such as intended use, observed risks, and study results as well as its corresponding appendices, including the study protocol.

10. Definitions

Actuation	The release of drug substance from the drug delivery device by a single activation (e.g., mechanical or breath).
Assay sensitivity	Ability of a clinical trial to distinguish an effective treatment from a less effective treatment or ineffective treatment.
Delivered/Emitted dose	Delivered dose is the quantity of drug substance that is available to the user, ex-device, on a per dose basis (i.e., released at the mouthpiece of the device).
Dose/Single dose	Amount of drug administered on a single occasion.
Fine particle dose	The quantity of drug substance with an aerodynamic particle size <5 µm on a per actuation of per dose basis. Used as a parameter for quality control.
Metered dose	Metered dose is the quantity of drug substance contained in the delivery device metering chamber.
Reference product	A product against which therapeutic equivalence is claimed.
Relative potency	The relative potency of the test product to the reference product is defined as the dose of the test product that produces the same biological response as one unit of the dose of the reference product (i.e., comparative outcomes for different doses).
Spacer/holding chamber	An add-on device for use with a pressurised metered dose inhaler (pMDI) consisting of a reservoir into which the aerosol is dispensed to aid inhalation.
Strength/dose	Strength is what is metered in the device for a single inhalation manoeuvre whereas a single dose may contain for example 2 puffs of a pMDI or 4 puffs of a pMDI. So, for example, for doses of 12µg and 24µg formoterol pMDI one and 2 puffs of the 12µg strength or two puffs of both the 6µg and 12µg strength might be used.
Single dose study	Single administration of each of the dose levels to be tested.

Product strength	Product strength may be either the delivered dose or the metered dose.
Pulmonary deposition	Amount of active substance deposited in the airways (mouth and throat excluded).
Therapeutic equivalence	The performance of the test and reference products are sufficiently comparable so that a clinically relevant difference between products with respect to efficacy and safety can be reliably excluded.

570

List of Abbreviations

APSD	Aerodynamic Particle Size Distribution
AUC	Area Under the Curve
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence Interval
C_{max}	Peak concentration
COPD	Chronic Obstructive Pulmonary Disease
DPI	Dry Powder Inhaler
FPD	Fine Particle Dose
GI	Gastrointestinal
ICH	International Conference on Harmonisation
IVIVC	*In vitro in vivo* correlation
MDI	Metered Dose Inhaler
OIP	Orally Inhaled Product
PD	Pharmacodynamic
PK	Pharmacokinetic
pMDI	Pressurised Metered Dose Inhaler
QWP	Quality Working Party
SmPC	Summary of Product Characteristics
TE	Therapeutic equivalence
t_{max}	Time to peak concentration