14 July 2025
CPMP/EWP/4151/00 Rev. 2
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 (MWP), Quality Working Party (QWP), Paediatric Committee/Formulation Working Party and Rheumatology and Immunology Working Party (RIWP)	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
Agreed by QWP	April 2025
Agreed by RIWP	April 2025
Adoption by CHMP	14 July 2025
Date for coming into effect	1 February 2026

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)".

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

Keywords	Guideline, inhalation, orally inhaled products, therapeutic equivalence, asthma, 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 that 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 for use in asthma and COPD.

The guideline focuses on abridged applications, but the principles described may be applicable for all 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 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;
- 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/EWP/QWP/1401/98 Rev.1/Corr**: Guideline on the investigation of bioequivalence;
- ICH M13A Guideline on bioequivalence for immediate-release solid oral dosage forms
- 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 should be carried out in accordance with Directive 2001/20/EC and Regulation EU 536/2024. Trials conducted outside the EU and intended to support a Marketing Authorisation Application in the EU/EEA have to be conducted according to the standards set out in Annex I of the community code, Directive 2001/83/EC.

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 are sufficiently comparable so that clinically relevant differences between products can be reliably excluded. The demonstration of TE between OIPs is based on a stepwise approach, where TE could be demonstrated *in vitro* in case all *in vitro* requirements are fulfilled, otherwise 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) are demonstrated in spite of some *in vitro*

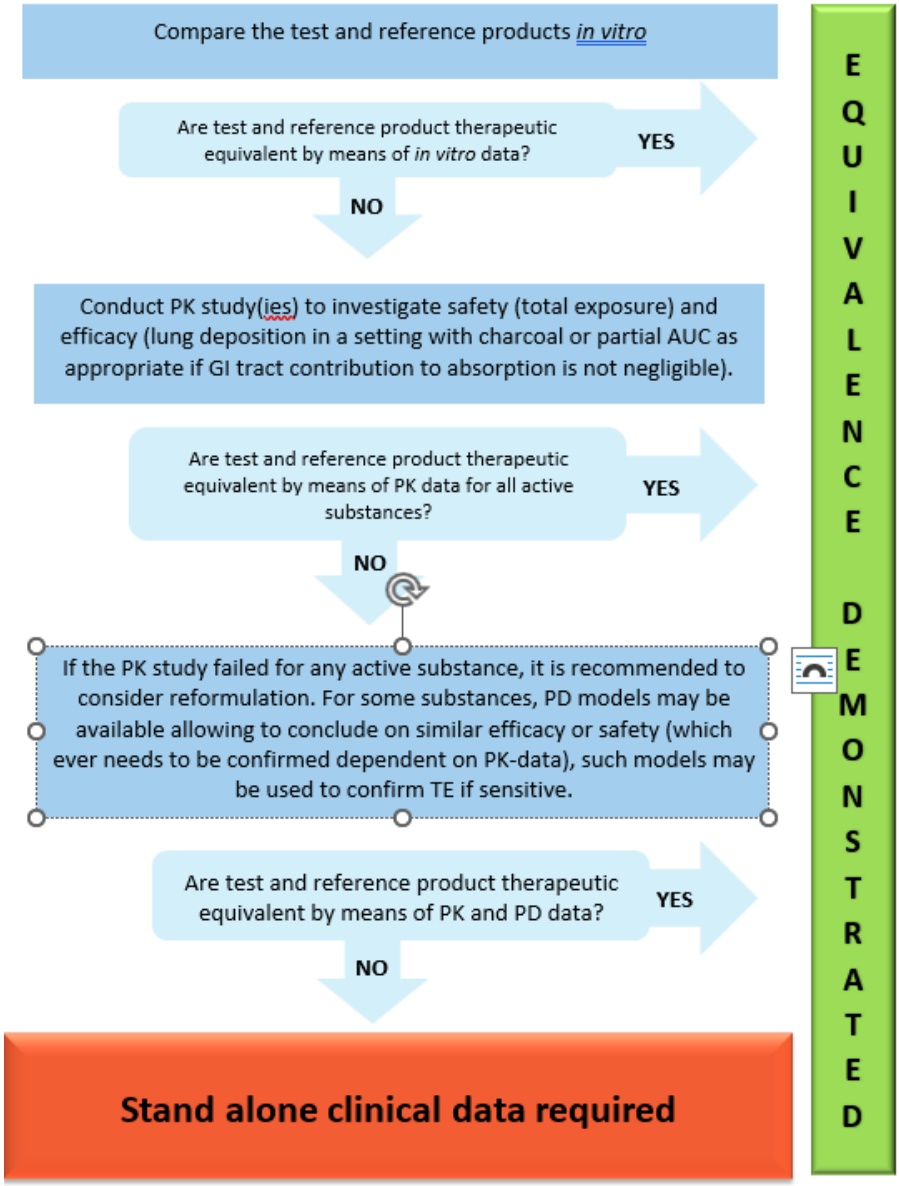
differences. It is generally not recommended to aim at demonstrating TE using pharmacodynamic or clinical endpoints as these are deemed insensitive (see Section 7).

The *in vitro* comparison process between the test and reference products is described in section 5. As stated above, the use of comparative *in vitro* data only 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, even in case some criteria are not fulfilled, and PK data are needed in addition (see Figure 1).

PK studies aim at evaluating pulmonary deposition and total systemic exposure as compared to the reference product, which is described in more detail in section 6. PK endpoints are considered valid surrogate markers to adequately predict similarity in the pattern and extent of deposition in the lungs and in the systemic exposure and, thereby, equivalence in both efficacy and safety. PK studies should generally be conducted in healthy adult volunteers. To assess pulmonary deposition, absorption of the active substance(s) from the gastrointestinal (GI) tract, if significant (contributing with ≥ 5% of the systemic exposure), 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 allow demonstrating TE regarding efficacy between the test product and the reference product based on PK studies, the test product should 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 based on PK studies it is sufficient to demonstrate that the systemic exposure is not higher with the test product when compared to the reference product.

Figure 1 Flow chart – demonstration of therapeutic equivalence of orally inhaled products



4.2. Additional considerations

4.2.1. Spacers

Spacers should always be considered with a pressurised metered dose inhaler (pMDI). Spacers are commonly used by children 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, studies generating data 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 recommended for use is to be replaced subsequently by an alternative spacer, appropriate data to demonstrate TE must be presented. Two *in vitro* studies need to be conducted with spacer. One study should be performed

comparing the aerodynamic particle size distribution (APSD) with a two seconds delay between firing and start of the air flow and with a relevant flow rate, e.g., 30 L/min. The delivered dose over tidal breathing should be compared in a separate study using a relevant breathing pattern, e.g., as described in Ph Eur 2.9.44 or other justified breathing pattern. The acceptance criteria as defined in section 5.1 should be used. In case TE is demonstrated using *in vitro* data for either the comparison with spacer or for the comparison 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 (e.g., based on *in vitro* data) that the spacer eliminates the fraction deposited in the throat. If a study with spacer and with charcoal blockade is waived, it should be noted that the spacer study without charcoal blockade is applicable to demonstrate TE regarding both efficacy and safety.

4.2.2. Products for nebulisation

This guideline applies also to products for nebulisation although it is acknowledged that the performance of these products is highly dependent on the nebuliser used. Data should be presented for at least one named nebuliser. The approach to demonstrate TE is as described in section 5.1 below. 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 similar.

4.2.3. Suprabioavailability

In cases of local suprabioavailability, i.e., if the test product displays an extent of pulmonary absorption appreciably larger than the reference product at the same delivered dose, 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 measures to minimize the risk should be provided.

4.2.4. Fixed combination medicinal products

For a fixed combination product of known active substances, TE should be demonstrated for each individual active substance. If 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 a study with charcoal if the charcoal administration was only necessary for the substance for which *in vitro* equivalence has already been demonstrated. Thus, in that case only a study without charcoal would be needed.

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 below, should be studied. If not all these *in vitro* criteria are 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 comparable to conclude on TE. The *in vitro* comparability exercise should be performed and evaluated based on a study protocol including methods of comparison and acceptance criteria. TE is sufficiently demonstrated if the test product fulfils all the following *in vitro* criteria as compared to the reference product:

1. The test product contains the same active substance as the reference product (i.e., 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 difference e.g., in crystalline structure and/or polymorphic form should not influence the performance of the product.
4. Any qualitative and/or quantitative difference in excipients must be adequately justified and it should be justified that it does not influence the relevant Critical Quality Attributes and/or any aspect of product performance other than those that are covered by the comparison of the APSD as described below (e.g., mouth/throat feel, taste, patients' compliance, or safety).
5. The 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. The inhalation devices 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 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. As the APSD of the reference product in the several test settings (different flow rates, different active substances in combination products, testing with and without spacer-use) will be different the need for grouping and the grouping strategy may apply differently for each conducted comparison of the APSD.
- At least four non-overlapping groups of stages or particle size fractions with defined cut-offs and not more than three 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 and pre-separator) and fine particle dose (FPD) should be evaluated as separate, additional groups, where the stages throat/induction port and pre-separator can be one group. The FPD should be represented by at least two 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%, assuming log-normal distribution of data (85-118%). In case of grouping, data on the corresponding individual stages should also be presented. 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. It is advised that feedback from the CHMP on such alternative approaches is sought through a scientific advice request.

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

Acknowledging that the number of *in vitro* comparisons of OIPs may be large, a comparison in one stage or group of stages not meeting the acceptance criteria might be acceptable in exceptional cases. Nevertheless, the number of batches and samples per batch investigated should be sufficient to minimise the risk for Type II-error[1]. 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 or units 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 (see section 5.2.3 below). In case *in vitro* comparison is only supportive for TE, and *in vivo* study(ies) are required, a smaller data set is considered sufficient, e.g., three batches and five inhalers or units of test respectively reference product.

5.2. Additional 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 drug product (formulation and device) 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 entire intended patient population. Patients may have impaired inspiratory ability as compared to healthy volunteers and thus differences in flow rate dependency may be of concern.

In principle, a study should be conducted in which the entire intended patient population and healthy volunteers inhale through the test product according to the instructions for use to obtain the pressure

[1] I. e. failing to reject the null hypothesis (ref and test are similar) when it's actually false (they are truly different). This is commonly caused if the statistical power of a test is too low, resulting in a "patient risk" accepting sameness when it is actually not true.

drops (and corresponding flow rates) that are representative of normal use of the product. Where applicable, in view of the indication, children with asthma and/or frail elderly should be included. The reference product should be added to this study if these parameters are not available from literature.

To investigate whether flow rate dependencies are sufficiently similar, the test and reference product should be compared at three points over the range of either pressure drops or flow rates reflective of healthy volunteers and the entire intended patient population, using the inhalation instruction as described in the respective SmPCs. This would in many cases correspond to 2 – 6 kPa.

If the difference in resistance to airflow between test and reference product is less than 15%, the comparison can be conducted at either three pressure drops (Option 1) or three flow rates (Option 2). If the difference in resistance to airflow between test and reference product is 15% or more, the comparison should be conducted at three pressure drops (Option 1).

For evaluation of similarity, the following graphs are expected:

a. The FPD at y-axis versus the square root of the pressure drop ($\sqrt{\Delta P}$) or the flow rate at x-axis.

b. The percentage of deposition (FPD) at y-axis versus the square root of the pressure drop ($\sqrt{\Delta P}$) or the flow rate at x-axis, where the FPD of the test and reference product at the highest studied pressure drop or flow rate should be set as 100%.

Similar flow rate dependencies can be concluded if the difference is not more than 15% between test and reference product at the lower two-point estimates in graph b.

If the flow rate dependency of the test product is lower than or similar to that of the reference product, PK studies in healthy volunteers are considered sufficiently representative for the entire intended patient population.

If the flow rate dependency of the test product is higher than that of the reference product or the relevant flow rates (covering healthy adults and the entire intended patient population) cannot be tested with the currently available cascade impactors, additional PK data, e.g., with trained healthy volunteers or patients, are required (see section 6.3.2 below).

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 the test and reference products should be demonstrated by *in vitro* testing.

If proportionality across all proposed product strengths is demonstrated with the test product, but not with the reference product, or vice versa, TE of the two products cannot be concluded 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, and 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 or groups of stages if justified (see section 5.1 above). The different strengths should be compared with a ±15% acceptance range in each stage or pre-specified groups of stages. 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 should 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 justified. For some inhalation products 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 five 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 calculated median of the observed reference product batches. 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 the 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 case 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 the test and reference products may indicate that the test and reference products are deposited in a different way in the lungs and absorbed at different absorption sites, and hence that there is a difference between the test and reference products that might 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).

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, salmeterol), a PK study with active charcoal can be performed to assess equivalence regarding efficacy. The charcoal blockade efficiency in terms of binding, charcoal dose and frequency of administration needs to be demonstrated (e.g., by *in vitro* binding studies and using a method that has been shown to be effective in the literature).

ii. Early partial AUC in a study without activated charcoal

In case the absorption of the drug in the lung is very rapid (e.g., median $t_{max} \leq 5$ min) and absorption occurs before the contribution of GI absorption is significant (e.g. salmeterol, glycopyrronium), so that it is possible to separate lung absorption from oral absorption, $AUC_{(0\text{-}30\ min)}$ is acceptable as a surrogate for efficacy and $AUC_{(0\text{-}t)}$ as a surrogate 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**) and ICH M13A. 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 that it can be avoided that C_{max} is observed in the

first sample post-dose. For example, formoterol and salmeterol have very rapid rates 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 $AUC_{(0-\infty)}$. If justified, a suitably truncated AUC can be used instead of $AUC_{(0-t)}$ for drugs with long terminal elimination half-life.

ii. Study population

Healthy adult volunteers generally demonstrate less variability in pharmacokinetic parameters than patients. In addition, patients may be less discriminatory since lung depositions are mostly central in case of bronchoconstriction. Therefore, 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 of the test product is similar to or lower than that of the reference product (see section 5.2.1), the study in healthy volunteers is sufficient.

If the flow rate dependency of the test product is higher than that of the reference product or cannot be determined with the currently available cascade impactors, TE cannot be concluded based on PK data in healthy volunteers only but additional PK data showing adequate exposure at the lowest relevant inspiratory effort are needed. This study may be conducted either in COPD patients with impaired inspiratory capacity or in healthy volunteers who are trained and monitored to inhale with low inspiratory effort, or by using a validated add-on device that increases airflow resistance. The use of activated charcoal blockade is not necessary. The objective of the study is to demonstrate that systemic exposure to the test product is not lower than that of the reference product. Specifically, the lower limit of the 90% CI for the ratio of the test to reference product for both AUC_{0-t} and C_{max} should not fall below the bioequivalence threshold of 80.00%. If systemic exposure for the test product is found to be lower than for the reference product, extrapolation from healthy volunteers to all patient categories would not be valid, and no conclusions about TE can 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 products 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 combination medicinal products, 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 the rare occasions when it has not been possible to find a representative batch, the development of an *in vitro in vivo* correlation (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 lower and upper tails of the distribution) representing the test product specifications are not inferior and not superior respectively 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 C_{max} and $AUC_{(0\text{-}t)}$ should be evaluated. In case an early partial AUC ($AUC_{(0\text{-}30\ min)}$) 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 products is contained within the acceptance interval of 80.00-125.00 for $AUC_{(0\text{-}t)}$ 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 $AUC_{(0\text{-}30\ min)}$ 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). Any deviations from these acceptance criteria must be thoroughly justified and can never be accepted for results below the lower limit of the acceptance range or when data on safety is generated from a study using different test and/or reference batches.

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 $AUC_{(0\text{-}t)}$ and C_{max} should not exceed the upper bioequivalence acceptance limit of 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.

6.4. *In vitro in vivo* correlation (IVIVC)

As discussed in section 6.3.2, 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 has not been established, normalisation will not be acceptable. The correlation should be shown for all active substances 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 active substance or no IVIVC can be 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, the 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

PK endpoints as described in this guideline are deemed to be the most sensitive to detect differences between the 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 on PD data without being demonstrated based on PK data, 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 to establish TE do not exist. In such cases, a full stand-alone clinical data package would need to be provided to support a MAA instead of a TE approach

Appropriate endpoints for TE from an efficacy viewpoint 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 the design.

Regardless of the objective 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 estimating assay sensitivity, it is essential to include at least one non-zero dose level besides the dose levels primarily 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 of 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 the statistical analysis allows calculating 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. To support

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

TE, it should 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, not previously 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 to use for a child 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 different risk/benefit relationship in the child compared with that seen in the adult.

In case 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 lowest approved age 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 the test and reference products.

PK data generated in adults may be used to support the demonstration of TE in adolescents (>12 years of age) without further justification. If the reference product is approved in children aged less than 12 years (whatever the lower limit of age), the Applicant is expected to provide a justification that the results of the PK study in adults can be extrapolated to the concerned paediatric population. For DPIs, a prerequisite for extrapolation of PK data from adults to children is the demonstration that the flow rate dependency of the test product is not higher than that of the reference product, or that an additional PK study has been provided demonstrating that the systemic exposure is not lower for the test product than for the reference product at a low inspiratory flow (see Section 6.3.2.)

9. Usability studies

For medicinal products for which 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 users 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. For such studies, 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 which both reference product-naïve and experienced users should be included. A minimum of 15 participants should be recruited in each distinct user group.

Participants selection for these studies should ensure representativeness of the intended users population incorporating general population trends (e.g., left handedness, elderly, patient with manual coordination difficulties, such as arthritic patients).

The study protocol should direct participants in simulating 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), unless a different study setting is justified. The exercise should include the unpacking of a new inhaler from the patient pack, simulated delivery of the first dose, as well as the intended storage of the inhaler. For pMDIs, the applicants should consider the use of placebo inhalers with propellant/excipients to assess the actuation force. Participants should be asked to simulate the delivery of further doses in order to

assess the user interface with the inhaler throughout its life. Areas of focus should allow ensuring that the user understands key features of the device.

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

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

10. Definitions

Abridged application	Application (generic, hybrid, biosimilar) submitted under Articles 10(1), 10(3) and 10(4) of Directive 2001/83/EC
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. One dose may consist of several actuations.
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/assessed.
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).
Single dose study	SA study involving a single administration of each dose level under investigation.
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 refers to the amount of active substance metered or delivered in a single actuation (e.g., 6 µg or 12 µg per puff of a pMDI). Dose, by contrast, refers to the total amount of active substance administered on one occasion, which may require multiple actuations. For example, to achieve a 12 µg dose of formoterol, one puff of a 12 µg strength inhaler or two puffs of a 6 µg strength inhaler may be used. For a 24 µg dose, this could be delivered as two puffs of the 12 µg strength, or four puffs of the 6 µg strength.

Therapeutic equivalence	The performance of the test and reference products is sufficiently comparable to ensure negligible impact on efficacy or safety.

11. 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