# E11A Pediatric Extrapolation

## Guidance for Industry

**U.S. Department of Health and Human Services**
**Food and Drug Administration**
**Center for Drug Evaluation and Research (CDER)**
**Center for Biologics Evaluation and Research (CBER)**

**December 2024**
**ICH-Efficacy**

# E11A Pediatric Extrapolation

## Guidance for Industry

*Additional copies are available from:*

*Office of Communications, Division of Drug Information*
*Center for Drug Evaluation and Research*
*Food and Drug Administration*
*10001 New Hampshire Ave., Hillandale Bldg., 4th Floor*
*Silver Spring, MD 20993-0002*
*Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353*
*Email: druginfo@fda.hhs.gov*
*https://www.fda.gov/drugs/guidance-compliance-regulatory-information/guidances-drugs*

*and/or*

*Office of Communication, Outreach and Development*
*Center for Biologics Evaluation and Research*
*Food and Drug Administration*
*10903 New Hampshire Ave., Bldg. 71, Room 3128*
*Silver Spring, MD 20993-0002*
*Phone: 800-835-4709 or 240-402-8010*
*Email: ocod@fda.hhs.gov*
*https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances*

**U.S. Department of Health and Human Services**
**Food and Drug Administration**
**Center for Drug Evaluation and Research (CDER)**
**Center for Biologics Evaluation and Research (CBER)**

**December 2024**
**ICH-Efficacy**

# FOREWORD

The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) has the mission of achieving greater regulatory harmonization worldwide to ensure that safe, effective, and high-quality medicines are developed, registered, and maintained in the most resource-efficient manner. By harmonizing the regulatory expectations in regions around the world, ICH guidelines have substantially reduced duplicative clinical studies, prevented unnecessary animal studies, standardized safety reporting and marketing application submissions, and contributed to many other improvements in the quality of global drug development and manufacturing and the products available to patients.

ICH is a consensus-driven process that involves technical experts from regulatory authorities and industry parties in detailed technical and science-based harmonization work that results in the development of ICH guidelines. The commitment to consistent adoption of these consensus-based guidelines by regulators around the globe is critical to realizing the benefits of safe, effective, and high-quality medicines for patients as well as for industry. As a Founding Regulatory Member of ICH, the Food and Drug Administration (FDA) plays a major role in the development of each of the ICH guidelines, which FDA then adopts and issues as guidance to industry.

*Contains Nonbinding Recommendations*

## TABLE OF CONTENTS


**I. INTRODUCTION (1) ..... 1**

**A. Objectives of the Guideline (1.1) ..... 1**

**B. Background (1.2) ..... 1**

**C. Scope (1.3) ..... 2**

**D. General Considerations (1.4) ..... 3**

**II. PEDIATRIC EXTRAPOLATION FRAMEWORK (2) ..... 4**

**III. PEDIATRIC EXTRAPOLATION CONCEPT (3) ..... 6**

**A. Disease (3.1) ..... 6**

**B. Drug Pharmacology (3.2) ..... 8**

**C. Response to Treatment (3.3) ..... 8**

**D. Safety Considerations (3.4) ..... 9**

*1. Extrapolation of Safety (3.4.1) ..... 10*
*2. Additional Safety Considerations (3.4.2) ..... 11*

**E. Sources and Types of Existing Data (3.5) ..... 11**

**F. Integration of Evidence and Development of the Pediatric Extrapolation Concept (3.6) 13**

**G. Establishment of the Pediatric Extrapolation Concept (3.7) ..... 15**

**IV. PEDIATRIC EXTRAPOLATION PLAN (4) ..... 15**

**A. General Considerations (4.1) ..... 16**

*1. Inclusion of Adolescents in Adult Trials (4.1.1) ..... 16*
*2. Modeling and Simulation Approaches (4.1.2) ..... 17*
*3. Dose Selection (4.1.3) ..... 18*
*4. Use of Dose Ranging Data (4.1.4) ..... 19*
*5. Use of Biomarkers (4.1.5) ..... 19*
*6. Establishing Relationships to Different Endpoints Between Reference and Target Populations (4.1.6) ..... 20*
*7. Safety Extrapolation Plan (4.1.7) ..... 20*

**B. Pediatric Extrapolation Plan Study Design Approaches (4.2) ..... 21**

*1. Exposure Matching Approach (4.2.1) ..... 21*
*2. PK/PD Approach (4.2.2) ..... 23*
*3. Efficacy Studies (4.2.3) ..... 24*
a. Single-arm efficacy studies (4.2.3.1) ..... 25
b. Externally controlled studies (4.2.3.2) ..... 25
c. Concurrent controlled efficacy studies (4.2.3.3) ..... 25
d. Incorporation of external data (4.2.3.4) ..... 26
e. Quantifying the impact of use of reference data (4.2.3.5) ..... 27
*4. Presentation and Justification for the Pediatric Trial (4.2.4) ..... 28*
*5. Analysis, Reporting, and Interpretation (4.2.5) ..... 28*
*6. Methods of Leveraging Reference Data in the Analysis of a Pediatric Trial (4.2.6) ..... 29*

*Contains Nonbinding Recommendations*

# E11A Pediatric Extrapolation Guidance for Industry[1]

This guidance represents the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA office responsible for this guidance as listed on the title page.

## I. INTRODUCTION (1)[2]

### A. Objectives of the Guideline (1.1)

The purpose of this guidance is to provide recommendations for, and promote international harmonization of, the use of pediatric extrapolation to support the development and authorization of pediatric medicines. Harmonization of the approaches to pediatric extrapolation should reduce the likelihood of substantial differences between regions. Importantly, harmonization should also reduce exposure of pediatric populations to unnecessary clinical trials and facilitate more timely access to pediatric medicines globally.

In general, FDA's guidance documents do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

### B. Background (1.2)

Regional guidances discussing pediatric extrapolation have previously been issued by various regulatory agencies. *Pediatric extrapolation* is defined in the International Council for Harmonisation (ICH) guidance for industry *E11(R1) Addendum: Clinical Investigation of Medicinal Products in the Pediatric Population* (April 2018) (ICH E11(R1)) as "an approach to providing evidence in support of effective and safe use of drugs in the pediatric population when it can be assumed that the course of the disease and the expected response to a medicinal product would be sufficiently similar in the pediatric [target] and reference (adult

[1] This guidance was developed within the Expert Working Group (*Efficacy*) of the International Council for Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) and has been subject to consultation by the regulatory parties, in accordance with the ICH process. This document has been endorsed by the ICH Assembly at *Step 4* of the ICH process, August 2024. At *Step 4* of the process, the final draft is recommended for adoption to the regulatory bodies of the ICH regions.

[2] The numbers in parentheses reflect the organizational breakdown of the document endorsed by the ICH Assembly at Step 4 of the ICH process, August 2024.

or other pediatric) population."[3] The reference population can include other pediatric age subsets. Pediatric extrapolation can extend what is known about the reference population (e.g., pharmacokinetics (PK)/dosing, efficacy, and safety) to the target population based on an assessment of the relevant similarities of disease,[4] drug pharmacology, and response to treatment between the two populations.[5]

Historically, extrapolation of safety generally was considered unacceptable. However, our understanding of similarities and differences between reference and target populations with respect to safety has evolved. As described in ICH E11(R1), the principle of using data generated in a reference population to define the scope and extent of data that should be collected in a target population can also apply to the generation of safety data (see section III.D (3.4)).

This guidance is intended to complement and expand on ICH E11(R1) to provide a more comprehensive framework for the use of pediatric extrapolation in optimizing pediatric drug development. The ICH guidance for industry *E11 Clinical Investigation of Medicinal Products in the Pediatric Population* (December 2000) (ICH E11) and ICH E11(R1) should be considered companion guidances. They provide guidance on the approach to safe, efficient, and ethical pediatric medicine development, including (1) when to initiate a pediatric program, (2) types of studies, (3) age categories, (4) age-appropriate formulations, and (5) ethical principles. ICH E11(R1) also encourages sponsors to engage regulators early in the development process to discuss approaches to optimize pediatric drug development, including the use of pediatric extrapolation and modeling and simulation (M&S). This guidance (ICH guidance for industry E11A *Pediatric Extrapolation*) is intended to aid sponsors and regulators on the degree to which pediatric extrapolation can be applied and the information that should be collected to address remaining uncertainties and gaps in knowledge supporting the safe and effective use of medicines in the pediatric population.

### C. Scope (1.3)

This guidance provides a framework for using extrapolation as a tool to support pediatric drug development. The framework describes an iterative process for understanding the existing information available, the gaps in information needed to inform development, and ways to generate additional information when needed. This guidance recommends approaches to assessing factors that influence the determination of similarity of disease, drug pharmacology, and response to treatment between a reference population and a pediatric target population. In addition, it discusses how the characteristics of the disease, drug pharmacology, and response to treatment may influence this determination.

This guidance discusses how the use of quantitative tools including M&S and other statistical approaches can be leveraged to fill in gaps in knowledge and/or reduce uncertainties. This guidance is not intended to provide a comprehensive listing of all diseases and/or situations

[3] See ICH E11(R1) at 5. We update guidances periodically. To make sure you have the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/regulatory-information/search-fda-guidance-documents.

[4] For the purposes of this document, the term *disease* includes both diseases and conditions. A condition may include a disease as well as being at risk for a disease.

[5] For the purposes of this guidance, unless otherwise specified, references to drugs or drug products include drugs approved under section 505 of the Federal Food, Drug, and Cosmetic Act (FD&C Act) (21 U.S.C. 355) and biological products licensed under section 351 of the Public Health Service (PHS) Act (42 U.S.C. 262) that are regulated as drugs

where extrapolation can play an important role in pediatric drug development. Rather it does explain how pediatric extrapolation can be applied practically to support the safety and efficacy of a product in pediatric populations. This guidance does not discuss other types of *extrapolation*, for example, the leveraging of foreign clinical data from one region for extrapolation to another region's population as a basis for registration of a medicine (see the ICH guidance for industry *E5 Ethnic Factors in the Accessibility of Foreign Clinical Data* (June 1998)). Although there are some quantitative strategies mentioned or explained within the guidance, it is not meant to be comprehensive. Some basic understanding of the role of quantitative approaches used in clinical trial development is expected.

### D. General Considerations (1.4)

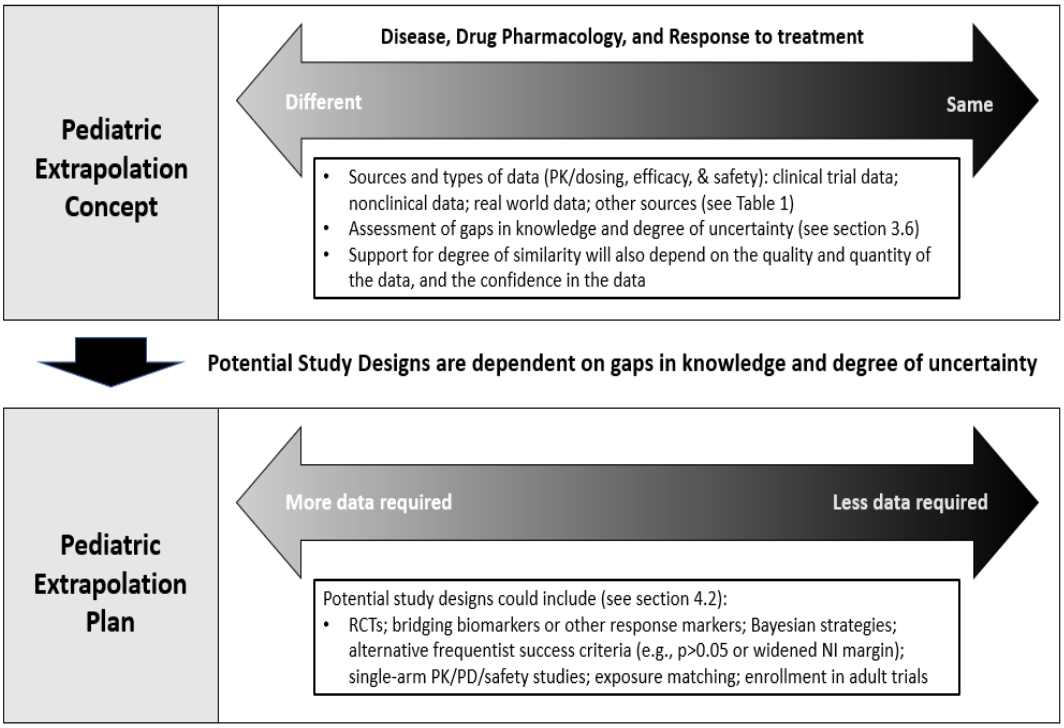
The use of pediatric extrapolation ensures that the pediatric population only participates in clinical trials when necessary to further the scientific understanding of the pediatric use of a medicinal product. As per ICH E11(R1), a sufficient prospect of clinical benefit is required to justify the risks of exposing the pediatric population to an investigational product. When pediatric studies are conducted as part of adult-driven drug development, including when these studies are required by regulatory authorities, the rationale for doing so can often implicitly assume a degree of similarity of the disease between the reference and target population. Thus, it may be appropriate for a pediatric program associated with an adult condition to incorporate some degree of pediatric extrapolation. While extrapolation to younger pediatric populations, particularly neonates, may be challenging due to rapid physiologic changes and organ maturation, the general principles in this pediatric extrapolation framework apply.

In the ICH E11(R1) definition of *pediatric extrapolation*, "sufficiently similar" might suggest a threshold that must be exceeded for pediatric extrapolation to be acceptable for regulatory consideration.[6] However, whether the disease and expected response to treatment can be considered sufficiently similar between a target and reference population is not simply a yes-or-no question. Therefore, this guidance does not use discrete categories (e.g., full, partial, none) to describe the different approaches to pediatric extrapolation, in favor of identifying the study designs which can address the remaining gaps in knowledge and uncertainties based on an assessment of the existing data. The use of extrapolation as discussed in this guidance reflects that a continuum of similarity/dissimilarity in disease, drug pharmacology, and response to treatment may exist between a reference and target population (Figure 1). The degree to which similarity is concluded will depend, in part, on a multidisciplinary assessment of the strength of the evidence, the confidence in the data reviewed, and the remaining gaps in knowledge. Importantly, the degree to which a reference and target population are determined to be similar as depicted in Figure 1 is not intended to illustrate a discrete *scale of similarity* but to provide a guide to understand this framework conceptually. Once the similarities between a reference and target population have been reviewed and remaining gaps in knowledge and uncertainties are identified, a pediatric extrapolation plan can be created by a multidisciplinary team. Options for study designs, methods, and analyses will depend on the gaps in knowledge and the level of uncertainty that needs to be resolved. The extrapolation plan should address these uncertainties, utilizing clinical judgement to establish the tolerable level of uncertainty that will be acceptable. The types of data and design of studies proposed in a pediatric extrapolation plan can range from exposure matching to randomized controlled trials (RCTs) (see Figure 1 and section IV.B (4.2)). Under

[6] See ICH E11(R1) at 5.

limited circumstances (e.g., when robust clinical information is available), there may be no requirement to collect additional PK/dosing, efficacy, and/or safety data in the target population.

**Figure 1: Pediatric Extrapolation as a Continuum**




PK = pharmacokinetics/pharmacokinetic; RCTs = randomized controlled trials; NI = noninferiority; PD = pharmacodynamic.
Section 3.6 = section III.F; section 4.2 = section IV.B.

## II. PEDIATRIC EXTRAPOLATION FRAMEWORK (2)

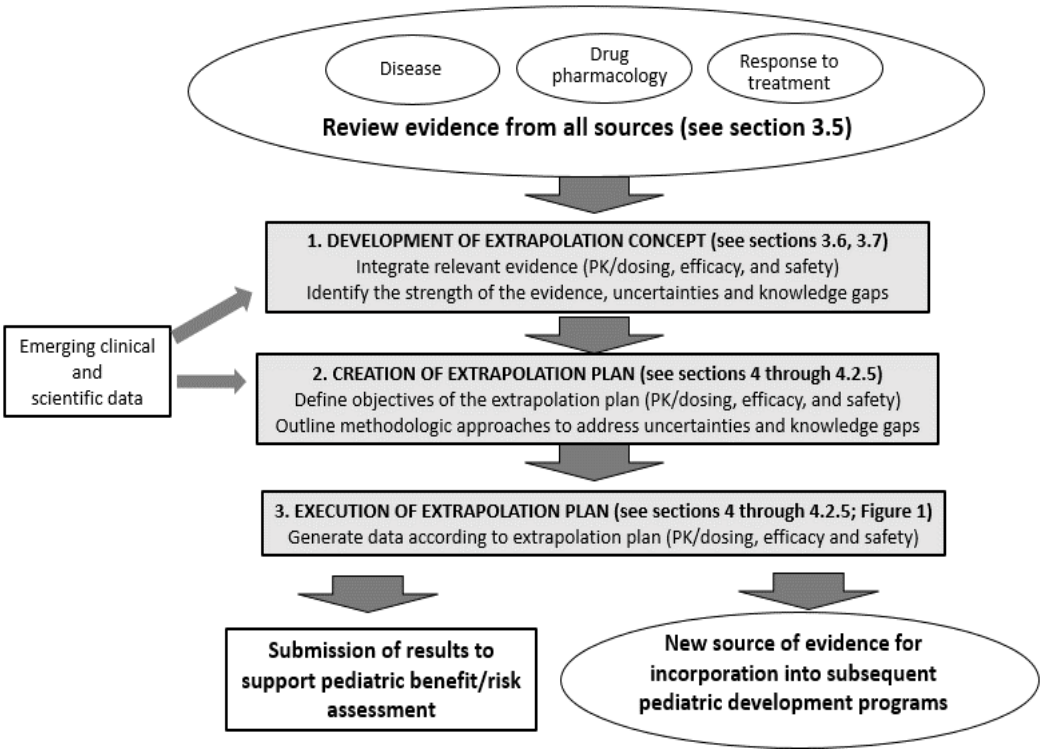
The extrapolation framework, as originally introduced in ICH E11(R1), is further refined in this guidance. The framework consists of three parts: development of a pediatric extrapolation concept, the creation, and then execution of a pediatric extrapolation plan (see Figure 2).

The first step is the development of a pediatric extrapolation concept, which serves as the justification for the pediatric extrapolation plan. The pediatric extrapolation concept evaluates what is known and unknown about the similarities and differences of a disease, drug pharmacology, and response to treatment between a reference and target population. The concept is developed through comprehensive and detailed review of existing information about the range of factors that define the disease, the drug pharmacology, and the clinical response to treatment across the reference and target populations. Factors that influence the effects of treatment in the reference and target populations should be identified and assessed. Once a review of the existing knowledge has been conducted, the data should be synthesized to develop the pediatric extrapolation concept. Methods to review and synthesize these data can include quantitative approaches such as M&S and statistical methods (see section III.F (3.6)). Synthesis of the data should be conducted to both understand the strength of the

evidence as well as to identify important gaps in knowledge which will inform what additional data may be required.

Once the pediatric extrapolation concept has been developed, the pediatric extrapolation plan should be created. This plan should include the objectives(s) and methodological approaches for the data that need to be generated to confirm assumptions made, address uncertainties and gaps in knowledge, and support benefit-risk assessment in the target population for the purpose of regulatory decision-making. In addition, there may be an evolution of the pediatric extrapolation concept based on emerging clinical and scientific data. In this case, rather than abandon an existing pediatric extrapolation plan based on a prior pediatric extrapolation concept, the pediatric extrapolation plan itself can be modified to reflect emerging scientific and clinical understanding.

**Figure 2: Pediatric Extrapolation Framework**




PK = pharmacokinetics.
Section 3.5 = section III.E; section 3.6 = section III.F; section 3.7 = section III.G; section 4 = section IV; section 4.2.5 = section IV.B.5.

The execution of the plan should also include an evaluation of the data generated to confirm any assumptions made and address uncertainties identified in the pediatric extrapolation concept.

The results should be submitted to support a pediatric risk-benefit assessment. These results should also be considered a new source of evidence for incorporation in subsequent pediatric development programs.

## III. PEDIATRIC EXTRAPOLATION CONCEPT (3)

Development of a pediatric extrapolation concept requires an understanding of the factors that can influence the disease, the drug pharmacology, the response to treatment, and the safety in both the reference and target populations.

### A. Disease (3.1)

The assessment of similarities and differences of the disease between a reference and target population is a key factor in developing the pediatric extrapolation concept. Although historically, pediatric extrapolation was often based on a binary determination of disease similarity (i.e., either yes or no), the understanding of similarities and differences in disease between a reference and target population has become more nuanced (see section I.D (1.4)). The evaluation of disease similarity is not intended to determine whether the disease in the reference and target populations is *exactly the same* but rather to determine the degree of similarities and/or dissimilarities of the disease. Even if there are differences in the disease, some similarities may be present that would still allow for the use of pediatric extrapolation.

It can also be possible to identify disease subgroups in both the reference and target populations that are sufficiently similar to support the use of pediatric extrapolation even if the disease in the overall population is not sufficiently similar. For example, many of the causes of adult heart failure are not similar to pediatric heart failure; however, heart failure due to dilated cardiomyopathy is similar between adult and pediatric populations, allowing for extrapolation from adult to pediatric patients with dilated cardiomyopathy.

To increase confidence in understanding the similarity of disease between the populations, evaluation of disease similarity should also attempt to determine the gaps in knowledge and uncertainties that exist in the evidence reviewed and identify what additional evidence is needed. Importantly, the evaluation of disease similarity is not a static or *one-time* exercise. As additional knowledge is gained, it should be incorporated into the evaluation of disease similarity in the pediatric extrapolation concept (see Figure 2). This evaluation may confirm or alter previous assumptions, resulting in either no impact or some impact to an extrapolation plan.

**Factors to consider in the evaluation of similarity of disease**

Assessment of disease similarity between a reference and target population should include a review of the following factors:

- **Pathophysiology of disease**

  Evaluation of the pathophysiology and etiology of the disease between the reference and target populations should be conducted. Collection of relevant information may include biochemical, genetic/epigenetic, cellular, tissue, organ system, and epidemiologic information that describes similarities and differences between the reference and target populations. Evaluation can also include a determination about whether differences in the clinical presentation of disease may depend upon the age of onset, age-dependent phenotypic expression, or other age-related differences. Evaluation of biomarkers that are common in the pathophysiology of the disease, including disease progression, if available, are often helpful in establishing

similarities in a disease between the reference and target populations. When possible, similarities in the outcome of untreated disease should also be evaluated.

- **Disease definition**

  Evaluation of disease definitions and diagnostic criteria between the reference and target populations should be conducted. When evaluating similarities and differences between reference and target populations, the following should be considered:

  — What are the manifestations or diagnostic criteria that define the disease?

  — How similar are the manifestations between the reference and target populations?

  — How are the manifestations measured?

  — Are there similar measurements used to define manifestations of the disease in the reference and target populations?

  — Are there subtypes (e.g., based on severity, genetics, molecular markers, etc.) of the disease that occur in the reference or target populations?

  — What are the similarities and differences in the subtypes of the disease in the reference and target population?

  — Are there other factors to consider (e.g., prognostic, predictive, genetic/epigenetic, psychosocial, etc.) that are needed to define the disease?

- **Course of disease**

  Evaluation of the similarities and differences in the course of disease between the reference and target populations should be conducted. In the evaluation, the following should be considered:

  — What are the similarities and differences of the clinical course of the disease between the reference and target populations? Are there differences in the course of the disease based on factors such as the age of onset of the disease?

  — Are there similar endpoints and/or biomarkers available that help to measure progression of disease in both the reference and target populations?

  — Are the short-term or long-term outcomes of the disease similar for the reference and target populations and can these outcomes be measured similarly?

  — What are the available treatments being used for both the reference and target populations?

  — What effect do these treatments have (e.g., timing of treatment relative to onset of disease and age of the patient, frequency of treatment, length of treatment) on the course of the disease in the reference and target populations?

Although the frequency, severity, or timing of the progression of the disease may differ between the reference and target populations, certain commonalities in the course of the disease may still allow for the use of pediatric extrapolation. If a treatment becomes available that changes the course of the disease in the reference population, but the treatment has not been approved in the target population, this should not necessarily lead to the conclusion that the course of the disease between the two populations is now different for the purposes of pediatric extrapolation.

### B. Drug Pharmacology (3.2)

Evaluation of the drug pharmacology for the purposes of pediatric extrapolation includes absorption, distribution, metabolism, and excretion (ADME) properties, pharmacodynamics (PD) (see section III.C (3.3)), and the mechanism of action (MOA) of the study drug.

Consideration should be given to the potential influence of intrinsic and extrinsic factors on ADME such as weight, body surface area, age, organ maturation, concomitant medications, and other relevant factors (e.g., protein binding, metabolic enzymes, transporters, renal function, or choice of dosage form). Differences in ADME properties can result in differences in PK parameters (e.g., clearance, volume of distribution) and resulting drug exposure. Exposure is a broad concept, ranging from measurement of the systemic (or other biological compartment) exposure of the drug (parent and/or metabolite(s)) at a single point in time (e.g., maximum or trough concentration) and/or exposure over a time interval (e.g., $AUC_{0\text{-}t}$ or average concentration).

When evaluating the PD and MOA of a drug, considerations should be given to the potential impact of maturation-related differences, for example, in expression level and sensitivity of the drug target(s) and, when applicable, potential downstream effectors. These differences may result in a different exposure-response (E-R) relationship for efficacy and safety between the reference and target populations. In addition, differences in secondary PD properties (i.e., off-target effects) of a drug may result in a different toxicity profile between the reference and target population.

### C. Response to Treatment (3.3)

To assess similarities and differences of response to treatment, a thorough review of available knowledge in both the reference and target populations should be conducted, including the response to the investigational drug, other drugs in the same class, and drugs in other classes, when used to treat the same disease. Similarly, data generated in other diseases for the drug, or drugs in the same class, can serve as a relevant source of knowledge when assessing similarities or differences of response to treatment. This assessment should include an evaluation of data on dose/exposure and response to treatment in the reference and target populations (see section IV.A.3 (4.1.3)).

**Factors to consider in the evaluation of similarity of response to treatment**

The degree of similarity of response to treatment between the reference and target populations can also support the degree of similarity of disease. A target (protein, receptor, mRNA, etc.) that is intrinsically associated with the disease in both the reference and target population leading to a similar therapeutic effect can support similarity of disease.

Assessment of similarity of response to treatment between a reference and target population should include a review of the relevant data on dose/exposure and response to treatment. The potential effect of developmental and maturational changes on the dose/exposure and clinical response should be a part of this evaluation. An understanding of the drug target and its role in normal development, disease pathology and expected response to treatment should be evaluated. For example, if a receptor does not exist in the first 6 months of life, no response to treatment would be expected for a drug only targeting this receptor in this age group. Factors that impact response that may differ between the reference and target populations (e.g., prior treatments, concomitant medications, comorbid disease, organ function, genetic makeup) should be evaluated to assess whether there is an impact on the extent to which pediatric extrapolation can be applied. In addition, understanding of the similarities and differences in the endpoints used to measure response can affect the overall assessment of similarity of response to treatment.

When evaluating the similarity of response, the following questions should be considered:

- Is there a similar measurement of the endpoint (e.g., clinical, biomarker, composite, etc.) used in both the reference and target populations?

- If the response endpoint or measurement of the endpoint is different in the reference and target populations, what is the relationship between the endpoints (e.g., clinical endpoint in the reference population in relation to a biomarker endpoint in the target population)?

- Are there factors (e.g., baseline severity of disease, prior treatments) that can affect both the exposure and the response?

When evaluating similarity of response to treatment, consideration should be given as to whether there are age/maturity-related factors (see sections III.A and III.B (3.1 and 3.2)) that could result in differences in the measured response between the target and reference populations. For many pediatric drug development programs, the primary endpoint(s) in the target pediatric population is/are different from that in the reference population. When this is the case, a comparison of one or more components of the primary endpoint(s) and/or secondary/exploratory endpoint(s) can be used to understand the relationship between the different endpoints. For example, if there is a biomarker that is correlated with an established clinical efficacy endpoint in a reference population, and if this biomarker is also correlated with clinical efficacy in a target pediatric population, such a *bridging biomarker* could support similarity of response to treatment (see sections IV.A.5 and IV.A.6 (4.1.5 and 4.1.6)).

### D. Safety Considerations (3.4)

Basic considerations for the development of an overall safety data collection and adverse event reporting plan are discussed in other ICH guidances (e.g., E2, E6, E9, E11, E11(R1)). This section describes specific considerations related to the extrapolation of safety as part of the overall development of the extrapolation concept.

### *1. Extrapolation of Safety (3.4.1)*

The principles underlying the appropriate use of data generated in a reference population(s) to define the scope and extent of efficacy data that need to be collected in a target population can also apply to the generation of safety data (see section IV.A.7 (4.1.7)). Extrapolation of safety data could be considered based on the available knowledge of the known and/or potential safety issues in the reference population that are relevant to the target pediatric population. Other relevant sources of information should be considered as part of this analysis (see section III.E (3.5)). These data should help increase certainty about the expected safety profile of a drug in a particular pediatric population and determine if additional gaps in knowledge need to be addressed in the pediatric program. Evaluation of the suitability and extent to which safety can be extrapolated should be included in the extrapolation concept and plan.

Similarities and differences in the safety profile between a reference and target population should be understood as a continuum. The source and amount of safety data to support the extrapolation of safety data to a target population should be considered early in drug development and planning. The reference population(s) can include pediatric and/or adult populations exposed to the same drug or class of drugs. Data can also be leveraged from reference populations that have been treated with different dosing regimens and/or for different diseases/indications. For example, the collection of safety data in adolescents, as defined in ICH E11 and ICH E11(R1), may provide a new source of evidence to support the safe use of a drug in younger patients. Enrollment of adolescents, in /or concurrent with the adult trials may also allow for earlier evaluation of safety for the adolescent population (see section IV.A.1 (4.1.1)).

When developing the safety extrapolation concept, the following questions should be considered:

- What is the age range of the target pediatric population to be studied as part of the safety extrapolation?
- What amount/quality of safety data are available from the reference population?
- Are there known on- or off-target effects of the investigational drug relevant to pediatric safety?
- Are data needed to account for age-specific, short-term, and longer term adverse effects in the target pediatric population, which may not have been identified in studies in the reference population?
- How does the expected treatment duration and treatment effect size in the reference population compare with the target pediatric population?
- How do the expected drug exposures in the reference and target pediatric populations compare? Does the exposure needed to target a specific PD effect or clinical response predict a specific toxicity in the target pediatric population?
- What information is already known from nonclinical sources (see Table 1) that can be leveraged to the target population?

- Are there other differences between the reference and target population that could limit the extrapolation of safety (e.g., a background therapy used in a target population that may potentiate a safety signal but is not used in the reference population, excipients in the formulation for the reference population)?

#### *2. Additional Safety Considerations (3.4.2)*

After an assessment of safety extrapolation has been made as part of development of the extrapolation concept, there may be a need to collect additional safety data over and above what has already been collected. This could be the case when there are remaining gaps and/or age-specific safety concerns in the target population (e.g., the effect of corticosteroids on reduction in growth velocity in prepubertal children with open epiphyseal growth plates). Consequently, it may be that longer term safety data should be collected in target pediatric populations post-approval.

Additional consideration as part of the pediatric extrapolation concept should be given to the collection of pediatric safety data in certain situations. Examples include:

- When the drug is a new molecular entity for a new class of drugs
- When there are known on- or off-target age-related safety concerns
- When there are significant safety findings noted in the reference population that would be of special importance in pediatrics
- When the drug has a narrow therapeutic index

Ultimately, the type, amount, and timing of the safety data that should be collected will depend on the gaps in knowledge identified as part of the pediatric extrapolation concept regarding safety in the target population(s). Moreover, clinical justification based on relevant available data should be the basis for establishing the size of the safety dataset; arbitrarily setting the size of the safety dataset is discouraged. Early discussion with regulatory authorities is recommended.

### E. Sources and Types of Existing Data (3.5)

Use of existing data should be fit-for-purpose (i.e., the context in which it was generated is applicable to the context in which it is intended to be used). It is important to consider both the quantity and quality of data to evaluate the similarities and differences between the reference and target populations. All relevant data should be used to establish the extrapolation concept and formulate the extrapolation plan. Such information may also include data from ongoing adult and pediatric development programs, or relevant data from terminated programs. Examples of the sources and types of data are included in Table 1 and are discussed further in this section. Table 1 is not intended to provide a comprehensive list as other sources and types of data may also be relevant.

**Table 1: Examples of Sources and Types of Data to Evaluate Disease, Drug Pharmacology, and Response to Treatment**

| Sources of Data | Types of Data |
|---|---|
| Clinical trial data | PK, PD, E-R, and clinical data in the same disease for the drug or drugs in the same class |
| | PK, PD, E-R, and clinical data in other related diseases for the drug or drugs in the same class |
| | PK, PD, E-R, and clinical data in the same disease for drugs in a different class |
| Nonclinical data | ADME data from animal models |
| | In silico, in vitro, and in vivo animal data (e.g., animal disease models, PK, PK/PD, MOA) |
| | Adult and juvenile animal toxicology data |
| Real-world data | Including but not limited to disease registries (regional, national, and international), electronic health records, health claims databases |
| Other sources | Including but not limited to systematic reviews or meta-analyses, relevant published literature |
| | Professional organization guidelines/clinical practice guidelines/consensus documents |
| | Published models/simulations (e.g., PK/PD, mechanistic) |
| | Expert opinion |
| | Standard of care/practice of medicine |

PK = pharmacokinetics/pharmacokinetic; PD = pharmacodynamics; E-R = exposure-response; ADME = absorption, distribution, metabolism, and excretion; MOA = mechanism of action.

## Clinical trial data

Clinical trial data (e.g., from controlled trials, prospective observational studies, PK, PK/PD and/or biomarker studies) in populations with the same disease or related diseases should be evaluated to understand similarities and differences between the reference and target populations. All relevant clinical trial data should be evaluated, including ongoing and completed studies, published or unpublished, whether results are positive or negative.

## Nonclinical data

Data from nonclinical sources such as in vivo, in vitro, and in silico models should be evaluated when available. These data may include PK, PD, and/or disease models. In general, when clinical data are available, data from animal models may be less relevant, but this is not always the case. In certain situations, disease similarity can be supported with **only** nonclinical data, especially when there is no ability to collect clinical data (e.g., anthrax or plague).

**Real-world Data**

The extent to which real-world data (RWD) can be used to support pediatric extrapolation, both the pediatric extrapolation concept and plan, is evolving. Therefore, the adequacy, relevance, and extent to which RWD can be used to support pediatric extrapolation should be discussed with regulatory authorities as appropriate. In the development of the pediatric extrapolation concept, a review of data from RWD sources, including but not limited to electronic health records, claims databases, and registries, should be considered.

**Other sources**

Expert opinions, including clinical practice guidelines developed by professional organizations, can be used to support the extrapolation concept. Published clinical practice guidelines from professional organizations are considered more informative than unpublished expert opinions. However, published guidelines and expert opinions can vary between regions based on differences in standard of care. Reliance on expert opinion or standard of care without an assessment of the strength of the evidence is generally not sufficient (see section III.F (3.6)).

In summary, the sources and types of data that are described above each have strengths and weaknesses. The confidence in the degree to which the sources and types of data support similarities between the reference and target populations require an assessment of the quantity and quality of data from each source as well as the context in which the data are being evaluated. A critical and multidisciplinary assessment of all the relevant data should be conducted to justify the use of the evidence to support the extrapolation concept.

### F. Integration of Evidence and Development of the Pediatric Extrapolation Concept (3.6)

The goal of the development of the pediatric extrapolation concept is not only to determine the acceptability to use pediatric extrapolation but also to describe assumptions made, detail any gaps in knowledge, and assess the impact of uncertainties in the available evidence. This section provides guidance on the review, synthesis, and presentation of information that should be included in a pediatric extrapolation concept.

**Integration of existing evidence**

Integration of existing evidence involves a comprehensive review to evaluate the similarities and differences of the disease, drug pharmacology, and response to treatment between a reference and target population (see Figure 2). Once the evidence is reviewed and integrated, the strength of the evidence is evaluated and gaps in the evidence are identified. Integration of the evidence should address the following questions:

- What is the body of evidence and what is the clinical relevance of the evidence?
- What are the strengths and the limitations of the evidence?
- How consistent are the findings across the sources and types of data?

- What inconsistencies exist in the data, and how do these inconsistencies affect assessment of similarity?

The answers to these questions will inform what additional information, if any, is recommended prior to establishing the extrapolation concept and/or what additional data should be collected in the extrapolation plan.

*Methodologies that can be used to integrate evidence*

Quantitative syntheses should be used to integrate the existing data (see Table 1) to develop the extrapolation concept. In order to adequately integrate the diverse types and sources of data, multiple quantitative approaches (empirical and mechanistic) could be employed and should be documented. Multiple approaches may be necessary to accurately leverage all data, given potential differences in key predictors for integration (intrinsic/extrinsic factors) and sources and types of variability (e.g., interindividual, intraindividual, and interoccasion). Systems biology/pharmacology models could be used to assess and predict disease biology, pathophysiology, and response to treatment. Population modeling (see section IV.A.2 (4.1.2)) could be used to inform estimation of key parameters and quantify sources of variability. Meta-analytic techniques could be used for synthesizing efficacy and safety data from multiple sources.

There are a variety of approaches available for quantitatively evaluating the similarity of disease and/or response to therapy in different populations (see section IV.A.2 (4.1.2)). Selection of an appropriate method will depend upon the data being evaluated for similarity assessment. For example, when using frequentist approaches, the evaluation of similarity of response between the reference and target populations can be informed by a comparison of point estimates and their associated confidence intervals. In many situations, it will be inappropriate to establish the degree of similarity purely based on overlapping confidence intervals. This evaluation should also take into account the precision of the point estimate and the magnitude of the difference between them. Bayesian hierarchical models could also be used to integrate and synthesize the available evidence. The manner in which uncertainty has been defined, specified, and otherwise accounted for in the evaluation, as well as any simulations used to assess similarity of disease and/or response, should be documented. In addition, any relevant assumptions with respect to the definition or expression of uncertainty should be specified.

Other exploratory analyses of the available data to assess similarity can also be considered. For example, if a trial conducted in a reference population has recruited across age groups, evaluation of the consistency of response in each age group can be considered. Approaches that can be used to evaluate the consistency of response across subgroups is described in other ICH guidances (see the ICH guidance for industry *E17 General Principles for Planning and Design of Multiregional Clinical Trials* (July 2018)).

When evaluating similarity of disease and/or response between reference and target populations, the available data may not permit definitive conclusions to be drawn given the inherent uncertainties in the data. As such, it is recommended that sponsors review the acceptability of the proposed approach with regulatory authorities, as appropriate.

**Knowledge gap identification**

Once the available evidence has been integrated, gaps in knowledge should be identified. It may be that some of the gaps in knowledge should be addressed before the pediatric extrapolation plan can be created based on emerging clinical and scientific data (see Figure 2). However, some gaps in knowledge do not necessarily preclude a pediatric extrapolation plan from being created. The pediatric extrapolation plan should address when and what data should be collected to fill these gaps in knowledge. Knowledge gap identification should address the following questions:

- What are the identified gaps in knowledge?

- Do these gaps in knowledge require additional data collection before the pediatric extrapolation plan can be created? If so, when and how will these data be collected?

- If these gaps in knowledge do not preclude creation of the pediatric extrapolation plan, when and how will these gaps in knowledge be addressed in the pediatric extrapolation plan?

### G. Establishment of the Pediatric Extrapolation Concept (3.7)

Establishment of the pediatric extrapolation concept should include a summary of the overall similarities and differences between the reference and target populations, the current knowledge gaps, uncertainties, and limitations of the data. This should include the following:

- An assessment of the evidence (i.e., overall strengths and weaknesses) of the similarities and differences between the reference and target population (disease, drug pharmacology, response to treatment). This should also include an assessment of the quantity and quality of evidence.

- An assessment of the available safety information and how this safety information affects the extrapolation concept.

- An assessment of the gaps in knowledge and how they affect the confidence and uncertainties in the extrapolation concept. In addition, this assessment should describe when and how the gaps in knowledge will be addressed.

## IV. PEDIATRIC EXTRAPOLATION PLAN (4)

Once a pediatric extrapolation concept has been established, the proposed study(ies) and/or data analyses (including model-based analyses) and their rationale should be detailed in an extrapolation plan. The design of the study(ies) and/or analyses should reflect the similarities and differences that have been identified between a reference and target population as well as the necessary information that needs to be collected to address the gaps in knowledge identified in the extrapolation concept.

## A. General Considerations (4.1)

As part of the development of a pediatric plan, there are some general considerations that may pertain to any study design selected (see section IV.B (4.2)). Some important general considerations include inclusion of adolescents in adult trials; use of model-informed approaches; considerations for dose selection; use of dose ranging data; use of biomarkers; considerations for endpoint selection when the endpoints differ between the reference and target population; and development of the safety plan. A discussion of each of these considerations is provided in this section. These issues should be considered as early as feasible in the development of a pediatric extrapolation plan.

### *1. Inclusion of Adolescents in Adult Trials (4.1.1)*

The enrollment of adolescents into adult clinical trials may hasten adolescent access to safe and effective treatments as well as accelerate the gathering of needed pediatric data. Historically, pediatric trials have not been initiated until after adult development has been completed and/or after the drug has been approved for adults. As a result, enrollment into pediatric trials may be slow due to the off-label pediatric use of the drug, further delaying broader pediatric and adolescent access to effective treatments. Inclusion of adolescents in some disease- and/or target-appropriate adult trials may address this problem. If the adolescent results are used to bridge the extrapolation of adult efficacy and/or safety to other pediatric populations, the similarity of disease, drug pharmacology, and response to treatment between these other pediatric populations and adolescents, and any gaps in knowledge, should be addressed.

The decision to include an adolescent cohort in an adult clinical trial assumes the disease, drug pharmacology, and response to treatment are sufficiently similar between the adolescent and adult patients. As such, the objective(s) of including adolescents and adults in a single trial should be framed within the context of the extrapolation concept. Additional data to inform adolescent dosing may not be necessary as adolescent and adult PK are generally similar. In such situations, specific consideration pertaining to the impact of lower body weight on dosing in adolescents should be carefully considered. In cases when there is a wide safety margin, higher exposures may be acceptable in the adolescents with lower body weight compared to adults when administered the same recommended adult fixed dose.

If the disease and response to treatment are sufficiently similar, the adolescent and adult populations can be combined into a single analysis of efficacy. The purpose and statistical methods for a separate analysis of the adolescent subgroup need to be carefully considered so that any identified differences or uncertainties are addressed. Such subgroup analyses should be interpreted cautiously; the strength of any conclusion about the extrapolation of efficacy (or lack thereof) based solely on exploratory subgroup analyses may be limited (see the ICH guidance for industry *E9 Statistical Principles for Clinical Trials* (September 1998)).

There may be ethical and operational challenges associated with including adolescents in an adult trial, such as: (1) different standards for the acceptable balance of risk and potential benefit; (2) whether adolescents should be exposed to a placebo control (which may be used more often in an adult trial); (3) the need for parental permission in addition to adolescent assent; (4) the use of the same primary endpoint and safety assessments in both the adolescent and adult population; and (5) the need for pediatric-specific study sites. If confronted with these challenges, different trial designs, or a separate adolescent trial run in

parallel to the adult trial can also be considered. In addition, there may be other challenges to the overall pediatric development program, such as pediatric investigators willingness to participate in a subsequent pediatric-only trial that may now exclude adolescents. Nevertheless, when the disease, drug pharmacology, and response to treatment are sufficiently similar between adolescent and adult participants, adolescents being included in an adult clinical trial or studied in a parallel trial is strongly recommended.

### *2. Modeling and Simulation Approaches (4.1.2)*

Modeling and simulation is an essential tool in pediatric drug development because it provides a means to address ethical constraints, data gaps, and logistical challenges, optimize dosing, and accelerate the development of safe and effective treatments for children. They also support evidence-based decision-making, leading to better outcomes for pediatric patients while reducing the risks associated with experimental trials (see ICH E11(R1)). Consistent with standard M&S practices, models should be developed and evaluated for adequacy and applicability for its intended purpose.

Modeling and simulation approaches are used in pediatric extrapolation, for example, to assess similarity of disease and response to treatment, examine and inform study design, derive dosing recommendations, test assumptions, and predict the effects of the drug in the target population. Quantification of relevant relationships (e.g., dose-exposure, E-R) provides an important foundation to conduct simulation in support of the dose selection. In addition, simulations of therapeutic window(s) associated with relevant PK or PK/PD endpoints can be explored prior to conducting a pediatric study. Modeling and simulation can be used to confirm the assumptions underlying the pediatric extrapolation concept after completion of the pediatric study. When simulations are used for regulatory decisions, it is important to provide information that the models are fit for simulation purposes and that model assumptions, input data quality, and the simulation set up are clearly reported. Typically, this information would be provided in the form of a modeling and simulation plan that the sponsor generates for internal documentation purposes, or a report suitable for interaction with regulators.

The availability of the various data sources dictates, in part, the modeling approach with more empirical approaches (e.g., individual PK/PD, population PK and PK/PD) reliant on data from reference population and mechanistic approaches (e.g., physiologically based pharmacokinetics (PBPK), quantitative systems pharmacology [QSP]) dependent on existing knowledge such as physicochemical, *in vitro,* and preclinical *in vivo* data. When using existing models (e.g., population PK, PBPK, population PK/PD models), the specific characteristics of the target population, such as relevant body size, organ maturation, and other relevant characteristics as needed, should be incorporated in the model. Depending on the available data and goals of the modeling, there are several techniques that can be used to incorporate information from the reference population in the analysis of the target population; for example, using models based on the reference population, analysis with pooled datasets, or Bayesian approaches (see section IV.B.3.d (4.2.3.4)) with prior distributions for model parameters. When selecting the appropriate technique, the advantages, disadvantages, and limitations of the selected technique should be considered carefully. For instance, when pooling data from both the target and reference populations, it is important to ensure an adequate representation of data from the target population. Failing to do so may result in parameter estimates being disproportionately influenced by the reference population and the inability to detect potential differences between the reference and the target population.

When making model-based assessments, the components of the model may have complex interrelationships (e.g., correlation of parameters and/or assumptions) that should be captured in the structure of the model along with any time dependencies. Model equations and assumptions underlying the model structure or dataset need to be clearly presented so that their relevance to the overall strategy, model predictions, and elements of uncertainty can be properly assessed. Assumption testing should be integrated into the model analysis.

It is important to distinguish between different sources of variability and uncertainties. For example, there is inherent variability in samples taken between individuals (i.e., between subject variability), which is a biological phenomenon and the magnitude of which can be directly supported by data. In addition, uncertainty about the model parameters can arise due to either incomplete data or to an incomplete understanding of biological or physiological processes (i.e., model assumptions). The different contributions of these sources of uncertainty should be addressed and justified. Procedures for estimating parameter uncertainty should be provided. Parameter uncertainty can often be reduced by incorporating more informative data. Given the extent of model assumptions and uncertainties, there should be multidisciplinary input to fully evaluate the M&S results and subsequent decision-making.

*3. Dose Selection (4.1.3)*

Evaluation and selection of an appropriate dose to be studied in the applicable pediatric subgroups is critical to achieve target exposures and responses. Before initiating pediatric studies, the available scientific information pertaining to the MOA of the drug, the pharmacokinetics of the drug (ADME), the effects of physiologic maturation of any organs and targets that are involved in the predicted exposures and responses to the drug and/or its active metabolites, and any additional relevant clinical data should be assessed (see section III.E (3.5)). As part of planning for dose selection, other considerations (e.g., safety, formulation, dosing regimen) should be incorporated.

Exposure-response relationships developed from data collected in a reference population can provide a strong pharmacological basis for justification of the exposure(s) ranges to be targeted. Subsequent simulations, incorporating relevant knowledge and available models, can be performed to inform dose selection (see section IV.A.2 (4.1.2)). The identification of safe and effective dose(s) in the reference population does not always require or result in the demonstration of an E-R relationship. As such, there is no requirement to establish an E-R relationship in the target pediatric population. Exposure matching may still be used in the absence of a demonstrable E-R relationship in the reference population and/or the inability to demonstrate similar E-R relationships in the reference and target populations. In situations where randomization of pediatric patients to subtherapeutic doses, or use of placebo, may be unethical and/or available safety data may not support evaluation of higher doses/exposures, generation of an E-R curve in the target pediatric population is not appropriate. In these circumstances, dose selection based on exposure matching is reasonable and pragmatic and is predicated on the expectation that a comparable response at the target drug exposure is likely to be achieved.

The aim of pediatric dose selection often is to target exposures similar to those known to be safe and efficacious in a reference population for further evaluation in a pediatric efficacy/safety study (see section IV.B.3 (4.2.3)). In order to confirm the selected pediatric dose, PK data are often necessary in the target population; however, a separate PK study may

not always be needed. PK data can be collected as part of the pediatric efficacy/safety studies with use of sparse PK strategies. When there is uncertainty about the proposed dose required to achieve a targeted exposure (e.g., developmental and maturational changes may impact PK), a lead-in PK assessment to evaluate the adequacy of the dose may be needed in the clinical study. A separate PK study should be considered in certain situations (e.g., drugs with narrow therapeutic range, non-linear PK, and/or potential differences in the effect of disease on the PK of the drug between the reference and target populations).

Alternatively, data in the reference population may be sufficient to predict doses in the target population using M&S (e.g., population PK, PBPK, or other M&S approaches). For example, PK data in the target pediatric population may not be required if there are exposure data on the investigational drug from a different pediatric population/indication of the same age and exposure range as proposed for the target population/indication. Additional PK data may not be necessary in the target population when observed exposure data are available in an adult reference population with the same disease, and the targeted exposure is within an observed exposure range in a different pediatric population of the same age with a different disease (s). However, these approaches rely on understanding the effect of disease on the PK of the drug. With adequate justification, there may be sufficient pediatric PK data such that M&S would be sufficient to establish an appropriate pediatric dose even if the observed exposures fall outside of the targeted range.

*4. Use of Dose Ranging Data (4.1.4)*

Dose ranging data may be needed as part of the pediatric extrapolation plan. Such circumstances may include when there is uncertainty in the disease similarity and/or response to treatment; when there are potential age-related differences in target expression; or when there is lack of correlation between systemic drug exposures and therapeutic response (e.g., locally acting drugs). Exposure-Response and/or dose-response (D-R) relationships can rely on a clinical endpoint or a biomarker response. Depending on the biomarker and the time course of the disease, dose ranging to achieve different degrees of biomarker/clinical response or an intrapatient dose titration to a target biomarker effect can be considered.

*5. Use of Biomarkers (4.1.5)*

Biomarkers can be used under different circumstances, such as a surrogate endpoint, in dose selection, and/or as a bridging biomarker (see section IV.A.6 (4.1.6)). When available, biomarkers that can be used to support both adult and pediatric development programs are desirable. Where relevant, it may be prudent to evaluate potential biomarkers to be used in a pediatric extrapolation plan as part of the adult development program. As an adjunct to the observed biomarker time course, a physiologic and/or mechanistic representation that describes the biomarker's relationship to disease progression and/or treatment effect is highly beneficial. M&S approaches can be useful for biomarker development strategy and choice of clinical endpoints in pediatric patients. Considerations for the use of biomarkers as part of a pediatric extrapolation plan should include the following:

- Use of a validated biomarker as a surrogate endpoint is recommended but not required, although use of a validated biomarker may require less justification.

- The choice of the biomarker endpoint should be supported by available data in the reference and target populations and justified in the extrapolation plan.

- A biomarker on the causal pathway that is correlated with clinical efficacy in the reference population is often acceptable and should be justified also with regard to its relevance to the target population.

- Models can be used to estimate the quantitative relationships between biomarkers and clinical efficacy (see section IV.A.6 (4.1.6)).

Methodological considerations with respect to the robustness and reliability of the proposed biomarker (e.g., the effect of missing data, sensitivity analyses, and departures from any assumptions) should be addressed.

*6. Establishing Relationships to Different Endpoints Between Reference and Target Populations (4.1.6)*

When developing a pediatric extrapolation plan, there may be differences related to the endpoint measurements that can be used to support efficacy between a reference and target population that should be addressed. For example, a clinically meaningful endpoint in a reference adult population such as 6-minute walk test distance is not suitable for pediatric patients who are pre-ambulatory. In such cases, evaluation of relationships between the endpoint used in a reference population and the candidate endpoint(s) in the target pediatric populations should be conducted. The following are considerations for identifying potential endpoints for use in a target population:

- Are there subcomponents of a composite endpoint that are similar between a reference and target population?

- Are there secondary endpoints in the reference population that could be used as a primary endpoint in the target population?

- Are there endpoint measurement scales that are similar between reference and target populations?

- Are there biomarkers that are correlated with clinical response endpoints in the reference population that are also correlated with clinical response in the target population (i.e., a bridging biomarker)?

The acceptability of the selected endpoint in the target population should be based on the strength of evidence available. When the proposed endpoint in the target population is different from the reference population, early discussion with regulatory authorities can be useful.

*7. Safety Extrapolation Plan (4.1.7)*

As described in section III.D.1 (3.4.1) (Extrapolation of Safety), the extrapolation concept should include a discussion of the extrapolation of safety in the target population and a thorough justification to support any conclusions about the acceptability to extrapolate safety information from the reference population to the target population. The approach to safety data collection described in the pediatric extrapolation plan should reflect the scientific question (s) that needs to be answered, the knowledge gaps identified, and the uncertainties

that are being addressed to support the safety of the drug in the target population. Even when extrapolation of safety is justified, there may be selective pediatric safety issues that should be addressed in the safety extrapolation plan (see the ICH guidance for industry *E19 A Selective Approach to Safety Data Collection in Specific Late-Stage Pre-Approval or Post-Approval Clinical Trials* (December 2022)). Additionally, if the safety margin is wide, it may be acceptable to target higher exposures than in adults. In the absence of a wide safety margin, it may be problematic to target exposures higher than adults and/or accept a higher degree of uncertainty in the predicted exposure. Under certain circumstances, no additional safety data will need to be collected beyond that which has already been collected in the target pediatric population as part of the efficacy extrapolation approach. If there is confidence that the available safety data collected are sufficient and address the relevant safety questions, there is no need to collect additional safety data in a pediatric preauthorization program.

Ultimately, the specific approach to safety extrapolation, including the potential need for pre- and post-marketing safety data collection, should be justified based on the safety extrapolation concept and discussed with regulatory authorities, as appropriate.

### B. Pediatric Extrapolation Plan Study Design Approaches (4.2)

The approach can range, for example, from matching effective and safe exposures in the reference population to generating concurrently controlled efficacy and safety data in the target population. The following approaches are discussed:

- Exposure matching approach (see section IV.B.1 (4.2.1))
- PK/PD approach (see section IV.B.2 (4.2.2))
- Efficacy studies (see section IV.B.3 (4.2.3)) including:
  - Single-arm studies
  - Externally controlled studies
  - Concurrent controlled studies

In addition, the design, timing, analysis, interpretation, and reporting of studies and/or analyses included in the pediatric extrapolation plan are discussed below.

Because an extrapolation concept can be considered as a continuum based on the gaps in knowledge and uncertainties, more than one study design may be appropriate to meet the objectives of the extrapolation plan. For example, there can be some overlap between the design of a single-arm PK/PD study and a single-arm uncontrolled study that relies on a clinical efficacy endpoint. In addition, an extrapolation plan can include a scenario that only requires evaluation of PK in the target population as the primary objective, but additional secondary clinical outcome measures can be included in order to increase confidence with the *PK-only* approach. Ultimately, the specific study designs used in any extrapolation plan should be justified based on the extrapolation concept and discussed with regulatory authorities as appropriate.

#### *1. Exposure Matching Approach (4.2.1)*

When there is strong evidence (1) to support similarity of disease between the reference and target population; and (2) that exposures in the reference population will provide similar response in the target population (e.g., infectious diseases, partial onset seizures), targeting

effective exposures in the reference population as the basis for pediatric extrapolation (i.e., exposure matching) may be reasonable. M&S strategies should be applied to support the initial dose selection in the PK study in the target population (see section IV.A.2 (4.1.2)). Allometric scaling should be used to account for weight-based changes in clearance and volume of distribution and maintain consistent exposures across various age/body weight groups. Models should also take into account other factors that may contribute to variability in exposures such as maturation (see section IV.A.2 (4.1.2)). In addition, model-informed dose selection should include an assessment of the feasibility and practicality of the dosing strategies. For example, fixed-dose combinations, dose volume limitations, and drug-device combination can influence the dosing strategy. Once PK data are obtained in the target population, the proposed dosing regimen should be evaluated through M&S techniques. If the proposed dose regimen does not achieve the intended exposure, then M&S can potentially be used to derive a modified pediatric dosing strategy that meets exposure matching criteria using available data without additional PK data collection in the target population.

**Target exposure metric and exposure range**

When the pediatric extrapolation strategy relies on matching exposures in the reference population, the target exposure metric(s), range, and acceptance criteria should be prespecified and should be defined in the context of the disease, MOA, treatment regimen, route of administration, and formulation. The chosen target exposure metric(s) should be associated with treatment response in the reference population and may be different for safety and efficacy. For example, AUC or $C_{min}$ may correlate with efficacy whereas $C_{max}$ may be more informative for safety. The target exposure range will subsequently be derived from established exposure-response relationships or observed data in the reference population. An adequate discussion and justification of the proposed metric(s) and range should be provided based on, but not limited to, the MOA and the metrics previously established in the E-R relationships in the reference population. It is often useful to present several exposure metrics. In cases where systemic exposure does not correlate with efficacy (e.g., most locally acting drugs), additional assessment of response might be needed.

**Sample size**

The proposed sample size for a pediatric PK study should be sufficient to meet the objectives of the study and can be based on quantitative methods (M&S and/or statistical approaches). Adequate representation of subgroups (e.g., body weight ranges, age ranges) should be considered and justified. The sample size justification and its feasibility in the targeted indication and age cohort(s) should include the following:

- The availability of patients in a specific body weight/age range;
- The availability of pediatric PK data from other disease populations;
- The adequacy of the sample size to demonstrate precision in key PK parameters in the pediatric population such as clearance and volume of distribution;
- The methodology(ies) used to determine the sample size.

Modeling and simulation techniques such as optimal design and/or clinical trial simulation should be conducted to inform the appropriate timing and number of PK samples. The timing

and number of samples collected should be aligned with clinical care whenever possible (see ICH E11(R1)).

**Analysis and reporting**

The analysis should evaluate whether matching the prespecified target exposure range in the reference population was achieved. Different presentations of the exposure data in the reference and target populations should be included. A single acceptance boundary for all drug products and drug classes (as compared to bioequivalence testing) will not provide a meaningful approach in the setting of pediatric extrapolation. An evaluation of exposure ratios and confidence intervals for the key exposure metrics (e.g., AUC, $C_{max}$, $C_{min}$) should be conducted. The chosen boundaries of the exposure ratios and the significance level of the confidence interval should reflect the context of the therapeutic range of the drug and the risk-benefit of the product for a given pediatric indication.

A descriptive comparison of observed exposures in the reference and target pediatric populations should be conducted. However, many pediatric programs rely on small sample sizes and/or sparse PK samples; therefore, model-based approaches in addition to descriptive comparison are generally preferred. In addition, a comparison of the model-derived exposure metrics integrating all relevant observed data should be performed when appropriate. Also, interindividual variability needs to be considered in establishing exposure similarity rather than comparing means alone. A simulation of the percent of participants at different age/weight ranges of the target population that lie within (or outside) a predefined exposure range may provide a more meaningful assessment of exposure similarity. Discrepancies between the observed/simulated data and target exposure range should be discussed.

In general, the most relevant factor to influence PK in pediatric patients is body weight. In addition to body weight, the developmental and maturational factors involved in drug disposition especially in the youngest pediatric patients (e.g., neonates and infants) should be considered. Relevant predefined exposure metrics should be presented graphically versus body weight and/or age on either a continuous or categorical scale. Relevant age and body weight ranges should be depicted in figures to allow for clear visualization of important covariates (e.g., dose(s), age, weight) as well as in tabular format. The reference range in the adult population (e.g., median and outer percentiles of the distribution of observed or simulated data) should also be presented graphically and in tabular format.

*2. PK/PD Approach (4.2.2)*

When exposure matching alone is insufficient to establish efficacy, PD biomarkers can be used as part of the extrapolation plan.

In order to rely on the use of dose/exposure to achieve a biomarker effect, it is important to have confidence that there is a relationship between the biomarker effect and efficacy in the reference population. Models could investigate the mechanistic basis for selected biomarkers, facilitate the analysis of biomarker data, and optimize the data collection needed to support and/or confirm the relationship between the biomarker and clinical efficacy in the reference population (see sections IV.A.5 and IV.A.6 (4.1.5 and 4.1.6)). A therapeutic range of the biomarker effect that provides a meaningful assessment of similarity of response between the reference and target populations should be defined in the extrapolation plan.

**Sample size**

In general, sample size considerations described in section IV.B.1 (4.2.1) apply. In addition, quantitative methods (M&S or statistical approaches) should be used to derive sample size for PK and/or biomarker endpoints. The sample size for the study can vary depending on the variability in PK and biomarkers. Consideration of the timing and number of data points per participant for both PK and biomarkers should determine the appropriate sampling.

**Analysis and reporting**

The data used in the analysis should be described, with a focus on the important elements relevant to the objectives of the analysis (e.g., the comparison between the biomarker effect in the target population and that in the reference population).

Results should be summarized with adequate graphical and tabular displays (e.g., illustrative plots for clinical interpretation). The clinical relevance of the results should be discussed, including the impact of any sensitivity analyses. The analysis and reporting should confirm a dose/exposure which results in the intended biomarker effect in the target population.

*3. Efficacy Studies (4.2.3)*

In many situations, efficacy data will be required to be generated. Efficacy studies can be performed to address different scientific questions including whether a novel treatment has a beneficial causal effect on efficacy but also whether the effect of treatment on outcome is similar in target and reference populations. Which is the most appropriate question will depend on the extrapolation concept and the gaps in knowledge and uncertainties identified. Different paradigms including frequentist and Bayesian may be applicable depending on the question of interest.

Regardless of the design chosen, as per the ICH guidance for industry *E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials* (May 2021) (ICH E9(R1)), a clinical study designed to demonstrate efficacy should have a prespecified estimand. Given potential differences between reference and target populations, certain attributes of the estimand may not align perfectly (e.g., dosing regimen, variable (endpoint), types of intercurrent events). When attributes differ, the estimand will always differ, and employing identical strategies for intercurrent events may not be feasible. However, for attributes consistent across populations, it is recommended to apply the same strategy for intercurrent events, and analysis method, whenever possible.

When incorporating external data into the analysis of the clinical trial, the estimand framework may also help in estimating the treatment effect of interest. To ensure valid comparisons and reliable estimation of treatment effects, potential biases, and confounders due to differences between the enrolled pediatric population and the historical control population (e.g., patient characteristics, disease progression), as well as differences in intercurrent event rates, should be carefully considered.

The design of the studies will be dependent on the gaps of knowledge identified in the extrapolation concept. One of the most important design decisions will be the choice of control arm. The options may include a single-arm trial with a formal comparison against an external control arm for which the data quality and relevance can be demonstrated, or a

randomized concurrent control arm. The choice will be influenced by the scientific question(s) identified in the pediatric extrapolation concept. For trials designed under the Bayesian paradigm, there are several additional options. The purpose of this document is not to provide prescriptive advice on model choice, but to provide important considerations when designing an efficacy study in the pediatric extrapolation plan.

a. Single-arm efficacy studies (4.2.3.1)

The use of single-arm PK studies in a target pediatric population in an extrapolation plan is discussed above (see sections IV.B.1 and IV.B.2 (4.2.1 and 4.2.2)). However, single-arm efficacy studies may be the most appropriate way of generating the required efficacy evidence. Situations where this could be the case include, but are not limited to, lack of a suitable control in the target population or when the accepted evidence for approval in the reference population is a single-arm trial. When designing such a study, how the primary efficacy objective would be evaluated should be defined using prespecified criteria such as a threshold for success, or prespecified precision. The threshold and precision should be established in the extrapolation concept and utilize clinical judgement to establish the level of uncertainty that will be tolerated.

The sample size of studies should be calculated so that there is adequate power to ensure the threshold for success is met, or that an estimate of sufficient precision is obtained. External data (e.g., published literature, available adequate RWD sources such as, electronic medical records, claims databases, or registries) can be used to contextualize the results with respect to current clinical practice, but without requiring a formal comparison of efficacy to external data.

b. Externally controlled studies (4.2.3.2)

It may be possible and appropriate in some circumstances to use external data as the formal comparator for a trial. This could be from the comparator arm in the reference population, relevant existing control arms from other RCTs in the target population, or RWD (e.g., results from observational studies) in the target population. Using external data such as from different pediatric populations  different diseases, or where different endpoints are used results in more uncertainty and may require additional justification.

As with any other study without randomized concurrent control, drawing causal inferences is more challenging. Since the data are compared directly with a data source external to the study, appropriate designs and statistical methods should be used to account for differences between the populations in order to minimize bias and confounding that can impact results. It is important to reflect that these studies would still be controlled, albeit with a nonrandomized control, which differs from the approach of just comparing to a threshold.

c. Concurrent controlled efficacy studies (4.2.3.3)

In some situations, the data generated to date and the outputs of the pediatric extrapolation concept are such that randomized controlled efficacy studies would be needed as part of the pediatric extrapolation plan to be able to draw benefit-risk conclusions. Based on the pediatric extrapolation concept, the need for controlled studies and the ability to extrapolate usually lead to study designs different than those that were required in the reference population. This will lead to a different relationship between the acceptable false positive

rate, false negative rate, and sample size that is not the same as it is in the reference population. When the sample size is limited, the relative importance of false positive and false negative error rates may be modified from convention (e.g., two-sided p-value less than 0.05) and what may have been deemed appropriate in the reference population. If a Bayesian design that uses an informative prior is used, strict control of the type I error rate is not possible. What is an acceptable trade-off between the risks of these two errors should be considered carefully on a case-by-case basis and discussed with regulatory agencies, as appropriate. When the sample size is determined based on a specific precision of the treatment effect, the error rates become less relevant.

It follows that extrapolation options may comprise many different design options that can be used to generate data, but not according to the conventional approach (e.g., an RCT analyzed in a frequentist framework requiring a two-sided p-value less than 0.05 for trial success). The extrapolation approach will usually result in a sample size smaller than one would expect for a standalone efficacy study. If the study is powered to meet a relaxed success criterion with a significance threshold greater than 0.05, such a modified frequentist approach should be justified in advance.

An alternative approach for designing active controlled trials may be to maintain the conventional type I error rate but widen the noninferiority (NI) margin usually used in *de novo* adult development, especially when the aim is not to demonstrate efficacy *per se* but to demonstrate that efficacy is in line with prior expectations based on the extrapolation concept. Alternatively, a wider confidence interval could be used.

Regardless of the approach used to demonstrate noninferiority, it will be important to ensure that the point estimate of the treatment effect does not raise concerns regarding inferiority.

d. Incorporation of external data (4.2.3.4)

When identifying which information will be incorporated into the analysis of the pediatric study, relevant data should be identified through an appropriate prespecified approach (e.g., systematic literature review using prespecified selection criteria). When possible, the sources of information to be leveraged should be agreed upon with regulatory authorities. However, it is possible that the external data themselves may not be available yet, for example, if generated from trials running in the reference population in parallel to the study in the target population or borrowed across age groups in the same study.

The types of information that could be leveraged in an analysis include individual patient data and/or aggregate data from other sources. Having access to individual patient data in the reference population enables comparison of the distribution of baseline prognostic factors with the target population. Potential differences between the study from which the reference data will be derived and the data generated in the target population should be adjusted and accounted for in the analysis. Bayesian and/or frequentist approaches can be used to combine data from the reference and target populations, weighting the contribution of the reference data based on an evaluation of similarity between reference and target populations. Borrowing approaches generally fall within one of two categories. The first category of methods evaluates similarity ahead of time, so that the degree of borrowing is prespecified and does not vary based on the observed data. Alternatively, dynamic borrowing approaches prespecify a model, which lets the degree of borrowing vary based on the consistency

observed in baseline characteristics and/or outcome data between the reference and target populations. Bayesian and frequentist versions of dynamic borrowing approaches exist.

e. Quantifying the impact of use of reference data (4.2.3.5)

Using additional data external to the trial using a frequentist approach leads to a clear understanding of how much data is being borrowed. This can be less clear using Bayesian methods. When using informative Bayesian priors, it is important to understand *a priori* how much available information is being incorporated into the analysis to support the interpretation of the pediatric trial. In particular, it is of relevance to know two separate pieces of information: how much of the information in the reference population do we expect to use in the exercise (i.e., the effective sample size (ESS) of the prior based on the reference data); and secondly, how much data will be generated in the target population relative to the prior ESS of the reference information being used. For some Bayesian models, there is a choice of methods for estimating the ESS. The ESS may be fixed or may vary depending on the observed data and the model chosen. In such cases, a range of ESSs should be documented. If the available information (based on reference data, or outputs from a M&S exercise) is summarized as a statistical distribution, then the ESS is a good way of describing how much information is being used.

If Bayesian approaches are used, different ways of using the prior information, for example by using a mixture prior (e.g., a prior composed of a mix of informative and weakly informative components) or power prior, will have a different ESS depending on assumptions made and the choice of parameters used in the construction of the prior. If such strategies are employed, sensitivity analyses evaluating the ESS under different values of these parameters will better help understand the design properties, especially in the case of prior-data conflict. Regardless of the approach used, the method of borrowing proposed should be prespecified and sensitivity analyses to understand the effect on operating characteristics of different amounts of borrowing will better help understand the design properties.

When there are known differences in disease between a reference and target population (e.g., disease severity), an extrapolation concept is still applicable when the differences can be quantified and adjusted for. When such differences preclude the use of the reference data as is, the data should be modeled to predict the efficacy in the target population more closely. In other situations, there may exist known differences in study design (e.g., the endpoint measured is different in the target population or the endpoint is measured at a different time) though the disease is considered to be similar to a degree that allows extrapolation. How the reference data are used in this situation would have to be considered on a case-by-case basis depending on the degree of similarity of disease, drug pharmacology, and response to treatment.

It can be possible to base a pediatric extrapolation plan on a biomarker, surrogate endpoint, or clinical endpoint as the primary endpoint in the target population, even if it is not the primary endpoint in the reference population (see ICH E11(R1) and section IV.A.6 (4.1.6)). In this scenario, an evaluation of the robustness of the correlation of the proposed endpoint to the primary efficacy endpoint in the reference population should be conducted. Where relevant, it may be prudent to initiate the evaluation of potential pediatric endpoints as part of the adult development program prior to their incorporation into the pediatric program.

*4. Presentation and Justification for the Pediatric Trial (4.2.4)*

Diagrams that represent the overall planned trial design are helpful, especially if the design is complex. This may be the case if, for example, there is an adaptive design or a trial with multiple stages evaluating different aspects of clinical development in each stage. When evaluating a trial design, determining what potential results will lead to a successful study based on predefined criteria can help to understand what magnitude of treatment effect would need to be observed for a trial to be declared a success. Tables or plots of different critical thresholds could be useful if there is uncertainty around the most appropriate threshold.

An evaluation of the study design should be conducted, including under scenarios inconsistent with planning assumptions such as where there is a prior-data conflict. This is especially important when Bayesian designs are used, including robust mixture priors. Regardless of design chosen (and whether a frequentist or Bayesian approach will be used), evaluations should establish operating characteristics of the design (e.g., false positive and false negative error rates), properties of the estimator (bias, variance), and properties of intervals (e.g., frequentist coverage of confidence or Bayesian credible intervals). Additionally, the results of a frequentist analysis of the data from the target population alone should always be provided.

*5. Analysis, Reporting, and Interpretation (4.2.5)*

Accurate analysis aligned with the prespecified estimand, thorough reporting, and clear interpretation of results are crucial to ensure reliable conclusions and informed decision-making. If a frequentist design is used, an alternative threshold to cross other than the standard two-sided significance level of 5 percent can be appropriate and should be justified and prespecified. A frequentist meta-analysis approach combining reference and target data could be conducted if the extrapolation concept supports that it is appropriate to formally analyze the data together.

If a Bayesian design is used, which explicitly leverages external data, there are many more choices to be made for the analysis. This analysis should be prespecified, although the prior may be updated as additional external data are generated. Visualizations to better understand the relationship between operating characteristics and underlying parameters and assumptions are helpful. Plots of posterior distributions may better contextualize the summary statistics derived from Bayesian analyses. If data external to the trial are incorporated into the analysis, the reporting should explicitly describe this and discuss how and when these data were originally generated, from what source(s) the sponsor acquired the data that go into the analysis, along with a justification as to why they are considered to be appropriate for inclusion.

Whether the data are sufficiently similar to enable combining to the extent proposed in the extrapolation plan, either from a Bayesian or frequentist perspective, depends upon the evidence generated in the target population. Ideally, the interpretation of a study is aided if the success criteria are described and agreed upon in advance with regulatory authorities, where appropriate. The criteria for success that could be used include a p-value, or if reference data are explicitly borrowed, Bayesian success criteria, such as credible intervals, excluding critical values, or the probability that one treatment is better than the other by at least a certain prespecified amount. More than one success criterion may be appropriate. For example, if an NI margin wider than would be accepted in adults is used, it is also possible to

specify the point estimate of treatment effect that would need to be demonstrated for noninferiority to be met for any given sample size and variance. This could help in demonstrating efficacy by providing additional reassurance of the expected treatment effect.

If the observed data in the study deviate from the observed reference data, this may limit the applicability of the pediatric extrapolation concept and the amount of data that may be considered reasonable to borrow. Nevertheless, if the data in the target population are substantially better than the reference population in terms of the point estimate of effect, but success criteria without borrowing has failed to be achieved due to a small sample size, it may be of interest to understand how much weight needs to be put on this reference data before a positive conclusion is drawn (i.e., using a tipping point analysis).

In some situations, one interpretation of data generated in the target population may be whether the data were consistent with what was expected based on the extrapolation concept. In such cases, study success criteria and the method of evaluation should reflect this study objective. Such interpretation of the data needs prior justification of the precision of the derived estimates and consequent sample size.

The more complex a statistical model, and the more parameters that need to be assumed, the greater the need for appropriate and wider ranging sensitivity analyses (see ICH E9(R1)). It is beneficial to discuss these sensitivity analyses in advance, and to investigate how robust the interpretation of the primary analysis might be to changes in these parameters. Such analyses should be carefully selected to investigate the assumptions made with the primary estimator and other limitations with the data.

#### *6. Methods of Leveraging Reference Data in the Analysis of a Pediatric Trial (4.2.6)*

The choice of reference data used in the analysis needs to be justified. When deciding on the method to use, simulation can be a useful tool to inform the choice of analysis strategy, with a view to optimizing the trade-off between bias, power (false negative), and type I error (false positive) rate. Various methods exist that aim to limit the borrowing if the data generated are not similar to the prior belief about them. As an example, one possible method amongst many is to use a robust mixture prior, such as a two-component mixture prior where one component is an informative prior based on the reference data and the second is a weakly informative prior i disregarding the reference evidence. The weakly informative component should be carefully chosen as specifying too large a variance may lead to substantial weight on extreme, clinically implausible parameter values which can impact the desired borrowing behavior. The prior weight attributed to the informative component of the mixture prior can be considered as the prior belief about the plausibility and acceptability of the extrapolation concept. The closer the value to 1, the more confidence there is. If small changes in the prespecified parameters such as the weighting parameter above, lead to large changes in the operating characteristics of the study, the method may not be sufficiently robust.

Sensitivity analyses are a useful tool for retrospectively assessing the robustness of conclusions to the strength of prior assumptions. One such example is a tipping point analysis where changes in the value of parameters specified in the Bayesian prior are assessed to see when there is a change in conclusion on efficacy.

When reference data are drawn from several different sources, such as adult RCTs, epidemiological studies, or registry data, the quality of data from the various sources may differ, and their relevance to the new pediatric trial may differ. In this case, careful consideration should be given to both the construction of the prior itself and the method used to include the data in the analysis.