EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

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

An agency of the European Union

2 December 2022
EMA/CHMP/ICH/205218/2022 Corr.*

Overview of comments received on ICH guideline E11A on pediatric extrapolation Step 2b EMA/CHMP/ICH/205218/2022

Please note that comments will be sent to the ICH E11A EWG for consideration in the context of Step 3 of the ICH process.

1. General comments – overview

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
European Hematology Association (EHA)	0	0		On behalf of the EHA Specialized Working Group on Pediatrics, we would like to extend our compliments for this well written and useful document. Our experts have nothing to add.	
ProPharma Group	0	0	0	In general, guidance to harmonise the approaches for paediatric extrapolation to support the development and authorisation of paediatric medicines, is appreciated. It is also appreciated to see the Section regarding Inclusion of Adolescents in Adult Trials (Section 5.2) to encourage drug developers to accelerate the gathering of paediatric data.	n/a
Takeda	0	0		This guideline is well written and provides important guidance about different study designs and considerations in extrapolation concepts/studies/analyses/reports. One challenge in pediatric clinical studies (either single arm or RCT) is slow enrolment. If it's not feasible to enrol the pre-planned sample size given the actual enrolment rate during the study conduct, can the Agency include some discussion on potential study design update in this guideline?	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Marko Ocokoljic, the European Society for Paediatric Oncology (SIOPE, or SIOP Europe)	0	0		SIOPE welcomes the development of the ICH guideline E11A on paediatric extrapolation, as a subsequent extension to previous consultations, that is sufficiently broad, highly relevant and supports the majority of circumstances where extrapolation could be of value, including in childhood cancer. Thus, we acknowledge and support this guideline which also addresses recommendations made by SIOPE in the reflection paper on the use of extrapolation in the development of medicines for paediatrics in 2018 (e.g. inclusion of MoA, extrapolation of PK/PD and clinical effectiveness relationship). Cancer medicine development is predominantly driven by adult cancer needs but many of the medicines in development have potential application in the paediatric population. There are many situations where data already generated by studies in the adult population could be used in an extrapolation concept to avoid unnecessary replication of studies, allowing the studies conducted in the paediatric (target) population to be appropriately focused on addressing the clinically relevant gaps in knowledge. Considering the developed ICH guideline, SIOPE specifically deems very important in childhood cancer: - The use of appropriate extrapolation and design of development plans in the new context of MoA driven paediatric medicine development (which is anticipated to be the hallmark of the paediatric evaluation of adult medicines for multiple disease settings); - crucial need for inclusion of adolescents in adult clinical trials in oncology (e.g. we are particularly in favour of the chapter on this topic at the end of the guideline, including the last sentence: "justify why not including adolescents in an adult trial"); - The essential role of academia in providing expertise and generating data, including standard clinical practice (e.g. an excellent example in paediatric oncology is the European Standards Clinical Practice (ESCP) Project, an on-going close collaboration between the European Reference Network for Paediatric Cancer (ERN PaedCan) and SIOP Europe's Clinical Trial Groups (ECTGs) – developing approved clinical recommendations reflecting current best practice for each common childhood cancer type).	
Pharmetheus AB	0	0		Clarification on extent of information that can be inhereted from adults to pediatrics, as well as, between pediatric age groups is needed. E.g. to what extent can informative priors based on the adult data be utilized in the analysis of the pediatric data; must the CL in 2-6 year olds be estimated with sufficient precision in a stand alone analysis, or can all pediatric data be considered in a joint analysis?	
Pharmetheus AB	0	0		Exposure response (E-R) is defined four times (rows 232, 244, 530, 536)	Define E-R once (row 232)
Pharmetheus AB	0	0		At several places, adolescents are not mentioned, eg line 359 "children and/or adults". This needs update to the text.	Change to children, adolescents, and/or adults
Agios	0	0		Consider harmonizing the terminology "external data", "reference data" and "source data"	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	0	0		**Specific examples:** consider including specific examples in the guideline directly, or in the associated training slides/materials, and in particular to better support extrapolation of safety; to provide clarification on extrapolation is a continuum, i.e., on what criteria are needed to fulfil or what aspects need to be discussed in an extrapolation concept so that to come to an overall successful conclusion. on extrapolation in the most challenging group,i.e., newborns and pre-term infants, for whom organ maturation is an important element to consider.	
EFPIA	0	0		**How an extrapolation concept could be translated into a pediatric development plan that employs extrapolation?** this is an important question to address since such a translation almost always needs a model to extrapolate efficacy from the reference to a target population. A model that we intend to use for extrapolation needs to be qualified for extrapolation. Nothing of this is really discussed in the current draft guidance. For example, such a model could be built and validated on adult data, linking exposure and baseline risk factors to clinical outcome. In order to qualify the model for extrapolation purposes, one could apply the model to the (available or to be generated) paediatric data, predict the clinical outcome (conditionally on the observed exposures and baseline risk factors in the paediatric data set), and compare the predicted clinical outcome with the observed clinical outcome. Such a comparison (if successful) would establish similarity and qualify the model for extrapolation. The model could then be used to estimate the treatment effect in the target paediatric population by applying it to a representative (for the paediatric population) set of exposures and baseline risk factors.	
EFPIA	0	0		**How to assess similarity between reference and target populations?** the question is key for an extrapolation concept, and it is closely linked to the model that is used for extrapolation. If there are no risk factors that need to be considered, then similar exposure should lead to similar efficacy. But if there are additional risk factors (like baseline severity) to be considered, then similar exposure will only lead to similar efficacy conditional that the risk factors are the same. A model that is used for extrapolation must therefore include all relevant risk factors. This type of discussion is almost completely missing in the guidance (only briefly mentioned in lines 842-847 or 868-876), whilst it is absolutely essential and should be the core of the guidance.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	0	0		**In addition,** the guideline focusses too much on statistical concepts that are driven by the understanding that one needs to compare an experimental treatment with a control (be it external or concurrent) in the target pediatric population. In these statistical concepts, the reference population serves to enrich this comparison, for example by using informative priors in Bayesian analyses or by using meta-analytic approaches for frequentist analyses. Using the reference population to enrich the comparison in the target population assumes that there is sufficient similarity. Whilst the guideline mentions that one should discuss this point in the extrapolation concept, it fails to provide guidance on how to assess similarity between reference and target populations. The concept of predictive distributions as a way to establish similarity is not truly present. This is critical in a context of extrapolation in order to account for small sample sizes appropriately.	
EFPIA	0	0		**Estimands:** throughout the guideline, reference is given to ICH E9(R1), but there is no mention of "estimand" concept relative to reference or target populations. This concept is very important in establishing the main questions of interest and the analytical methods tasked with answering them. It is suggested to mention that the estimands concept as described in ICH E9 (R1) should be used as relevant, depending upon the extrapolation strategy.	
EFPIA	0	0		**Frequentist or Bayesian approaches**: the guideline should treat equally these 2 approaches in terms of expectations. Many aspects are relevant to both approaches while one might be more appropriate in some circumstances versus the other. The considerations to Frequentist or Bayesian approaches are otherwise spread and not equally treated. For instance, the need to define decision criteria, to evaluate the operational characteristics, to foresee sensitivity analyses,....are valid either ways with some nuances linked to the specificities of each statistical framework	
EFPIA	0	0		**Terminology:** "extrapolation of data" is mentioned throughout the guideline including in the safety section. This wording is misleading since data itself can be analyzed or interpreted but not extrapolated. One is extrapolating the treatment effect, or the outcome measures, from a reference population to a target (pediatric) population. Our recommendation is to replace "extrapolation of data" by "extrapolation of treatment effect" or "extrapolation of (safety/efficacy) outcome".	
EFPIA	0	0		**Extrapolation in pediatric populations between different race/ethnicity**: this is missing from the guideline (e.g., from Caucasian pediatric population to Chinese/Japanese pediatric population). Inclusion of population factors within the model should be included within baseline risk-factors. This also relates to having a comprehensive MIDD approach which is missing from the guidance. Although this topic should belong to ICH E5 topic, there is no explicit guidance in E5 either, therefore it's important to be included here.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	0	0		**Companion diagnostics (CDx):** there are no considerations in the guideline for CDx strategies in pediatric indications. Development of CDx in small pediatric indications can be challenging and regulatory guidance that provides flexibility would be helpful. Moreover, there are also no consideration on medical devices, while reference could be made to the 2016 FDA guidance on Leveraging Existing Clinical Data for Extrapolation to Pediatric Uses of Medical Devices: – medical devices for extrapolation in paediatrics: ttps://www.fda.gov/media/91889/download).	
EFPIA	0	0		**Glossary of terms:** consider including one, as the guideline uses many terms, but their meaning in the context of the guideline is unclear. For example, the following could be considered:	
EFPIA	0	0		The term **"similarity of response"** needs to be discussed and defined better. Does it refer to the outcome under the experimental treatment, or does it refer to the treatment effect relative to a control group (i.e., to the difference between experimental treatment to a control). I think it should be the former, but this should be made clear.	
EFPIA	0	0		what is **"uncertainty"**? Is it "lack of precision" (due to small sample size) or does it refer to "questions about interpretation of a result" (i.e., potential bias)?	
EFPIA	0	0		Does **"strength of evidence"** (line 432) mean "low uncertainty", "absence of gap in knowledge", or something else? On the same note, use the same term if you mean the same thing.	
EFPIA	0	0		Sometimes the guideline uses the term **"source"** population, sometimes **"reference"** population. Unless these terms describe different things, use just one of the terms. If they do describe different things, provide a clear definition for each so that the difference becomes clear. A similar comment applies for "exposure-response" and "PK/PD" relationship.	
EFPIA	0	0		Throughout the document, the term **"synthesis of data"** is used but meaning is unclear; please give a definition for "synthesis of data".	
EFPIA	0	0		The guideline seems to use **disease and condition** interchangeably. For example, in Table 1, with "same condition", do we mean "same indication"? It is suggested that clarification of the difference between these terms be provided, if any. Otherwise only one of these terms should be used throughout.	
EFPIA	0	0		The guideline uses the terms **"drug pharmacology**" compared to "drug (pharmacology)" throughout the document. We suggest using the term "drug pharmacology" to ensure consistency and clarity. If "drug (pharmacology)" means something different from "drug pharmacology," we suggest clarifying the meaning of both terms.	
EFPIA	0	0		**"Data extrapolation" vs. "findings extrapolation":** data itself is typically not extrapolated, but findings are. It is suggested for example, to replace "extrapolation of safety data" with "extrapolation of safety findings" (Line 348) and extrapolation of data" with extrapolation of findings" (Line 47).	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation

2. Specific comments on text

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
ISCT	1	1017	All	The document makes the point that it is a guideline and is not prescriptive (e.g., in line 54 it says it 'is not meant to be a comprehensive instruction guide.'). Paediatric extrapolation is very complex and associated risks are high. It requires a high level of multiple expertises. Therefore it should emphasise that it is an approach to how to go about evaluating potential for paediatric extrapolation and undertaking it where justified.	
Lundbeck	1	1	1	General comment: The term "extrapolation" is somewhat confusing, as it is usually used to conclude on something beyond what has been observed, based on known results, without gathering new data. In this document typically some data will be gathered on the target population. It is clearly explained in the document that the amount of extrapolation should be seen as a continuum, but it seems to contrast with the actual meaning of the word "extrapolation".	
Lundbeck	1	1	1	General comment: the structure of the document is sometimes somewhat confusing. Sometimes topics are discussed in a specific chapter that seem to apply to other chapters as well, but because they are mentioned in that specific chapter it becomes unclear whether these considerations apply more generally as well. See for example comments from rows 24 and 28.	
ISCT	4	5	1,1	Suggest to caveat lines 4 - 6 ("The purpose of this guideline is to provide recommendations for, and promote international harmonization of, the use of pediatric extrapolation to support the development and authorization of pediatric medicines.")	Suggest change to words similar to: 'The purpose of this guideline is to provide recommendations for, and promote international harmonization of, APPROACHES TO HOW pediatric extrapolation MAY BE USED WHEN JUSTIFIED to support the development and authorization of pediatric medicines."
German Pharmaceutical Industry Association	6	7	1	Harmonization between regions is strongly supported.	
AESGP	6	7	1	Harmonization between regions is strongly supported.	
GPT	8	9		GPT supports the avoidance of unnecessary exposure of minors of age to clinical trials as a crucial value. This is due to the a priori experimental character of clinical trials.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
German Pharmaceutical Industry Association	8	9	1	The avoidance of unnecessary exposure of minors to clinical trials is supported as an important value. Also should a more timely access to pediatric medicines be in focus.	
Agios	8	9	1,1	Important to specifically call out that one of the intended objectives is to reduce exposure of pediatric populations not only to unnecessary clinincal trials but also to non-active comparators in a clinical trial when this could be avoided in certain circumstances by leveraging extrapolation.	Importantly, harmonization should also reduce exposure of pediatric populations to unnecessary clinical trials, and in some cases to non-active comparators, and facilitate more timely access to pediatric medicines globally.
AESGP	8	9	1	The avoidance of unnecessary exposure of minors to clinical trials is supported as an important value. Also should a more timely access to pediatric medicines be in focus.	
GPT	12	20		The similarity concept is deemed as adequate and helpful to avoid unnecessary exposure of minors of age into clinical trials. This holds particularly true for minors of age close to adulthood. However, a stepwise approach may extend this.	We propose implementing a stepwise approach from minors of age, i.e. starting with the older teenager group to the younger ones and so forth.
German Pharmaceutical Industry Association	12	20	1	The concept of similarity seems to be helpful to avoid unnecessary exposure of minors to clinical trials. This is especially true for minors close to adulthood. However, a stepwise approach should be implenented	A stepwise approach from minors of age should be implemented. I.e. starting with the older teenager group to the younger ones, and so on.
AESGP	12	20	1	The concept of similarity seems to be helpful to avoid unnecessary exposure of minors to clinical trials. This is especially true for minors close to adulthood.	
EFPIA	18	18	1,2	The text states that safety can now also be extrapolated between populations. Clarification is requested as to whether this applies only to target/MoA-related safety.	Please clarify the statement.
Koop Phyto	19	19	1.2	not only assessment of the relevant similarities of disease and response to therapy, but also assessment of the relevant similarities of adverse reactions/undesired events is essential.	see next line
Koop Phyto	19	20	1.2	Subsuming side effects and adverse events as a special case of treatment response does not do justice to their importance.	… assessment of the relevant similarities of disease, response to therapy, adverse reactions and undesired events of the two populations.
German Pharmaceutical Industry Association	22	26	1	The extension of the principle of using data generated in a reference population to safety data is strongly supported.	
AESGP	22	26	1	The extension of the principle of using data generated in a reference population to safety data is strongly supported.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	22	26		It is welcome to see extrapolation being applicable to safety in principle. Can ICH provide a few examples regarding safety extrapolation?	
ISCT	26	26	1,2	Suggest to caveat line 26 because ICH 11 (R1) caveats the approach to leveraging safety data from the reference population to the paediatric population where it says: "When efficacy in the pediatric population can be extrapolated from data obtained in the reference populations, leveraging of safety data from the reference to the pediatric population may be utilized; however, additional pediatric safety data are usually required, as existing data may only provide some information about potential safety concerns related to the use of a drug in the pediatric population [See ICH E11 (2000) Section 2.4]." (per para 2 on page 8 in ICH E11 (R1))	Suggest to caveat line 26 by words similar to the following: '...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 MAY also apply to the generation of safety data SUBJECT TO MEETING CERTAIN UNIQUE REQUIREMENTS IN THE TARGET POPULATION (see section 3.5 THIS DOCUMENT).'
ISCT	29	29	1,2	Suggest to modify line 29 as indicated in next cell because E11A (2b) is a guideline in a complex area	Suggest: '…comprehensive framework TO GUIDE THE use of paediatric extrapolation…..'
ISCT	30	30	1,2	Roadmap' seems too prescriptive for something as complex as paediatric extrapolation, and drug developers and regulators are responsible for assessing complex information in their risk assessments.	Suggest: 'This guideline IS INTENDED TO aid to drug developers and regulators in ASSESSING THE degree to which…'
GPT	35	43		In the entire extrapolation concept, combinations products are not specifically mentioned.	We recommend including combination products in general in the overall concept of extrapolation.
GPT	35	43		The iterative approach is deemed as suitable for the purpose to limit and tailor trial designs that are adequate for the pediatric population.	
German Pharmaceutical Industry Association	35	43	1	The iterative process for understanding the available information in oder to identify gaps seems appropriate for the purpose of designing adequate trials for pediatric populations.	
AESGP	35	43	1	The iterative process for understanding the available information in oder to identify gaps seems appropriate for the purpose of designing adequate trials for pediatric populations.	
EFPIA	37	39		This is the core described in the sentence, but the sentence seems not appropriately structured. Suggest rewording.	….understanding the existing informationa available, identification of the gaps in information needed to inform development and ways to fill the gaps and potentially generate additional information when needed to support….

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Takeda	42	43	1,3	Recommend to also consider the differences in physiology in the pediatric population and their ability to clear and/or metabolize drugs particularly in the much younger subsets of the pediatric population	
Koop Phyto	43	43	1.3	Extrapolation cannot only be used to develop new drugs but also to strengthen the application of already existing products that are well-established and medicinally applied since many decades or even centuries in the paediatric population. Many of these products with such long-standing use have a wide-spread off-label-use (e.g. not authorised for one or more paediatric age groups) in paediatric real-life, however with few documentation. The already existing experiences with such medicinal products should not be lost. Data collection and evaluation should therefore also include the available experience in this field.	Please add: In addition, the guideline discusses the use of already available scientific knowledge on well established drugs in order to support the extrapolation of the use of existing medicines for their use children.
German Pharmaceutical Industry Association	45	50	1	It is very positive that this guideline focusses on practical support and does not try to be comprehensive, reflecting the ongoing development of statistical and quantitative tools.	
AESGP	45	50	1	It is very positive that this guideline focusses on practical support and does not try to be comprehensive, reflecting the ongoing development of statistical and quantitative tools.	
EFPIA	45	46	1,3	It is not only a question of tools but also data; also it could be reformulated slightly	The guideline discusses how the use of statistical and other quantitative tools (e.g., such as modeling and simulation) may be leveraged to fill in gaps in knowledge.
ISCT	46	46	1,3	Suggest to caveat	Suggest 'MAY' instead of 'CAN'
GPT	46	48		It is adequate that this guideline is not meant to be comprehensive, as this reflects the multiple options including scientific or society's evolution about the topic.	We propose to indicate that progress in methodology or society values may change the approach to particular study designs.
ISCT	47	47	1,3	Suggest to caveat	Suggest 'MAY' instead of 'CAN'
EFPIA	47	47	1,3	The terminology "extrapolation of data" was first mentioned here and elsewhere in the document (in the safety section). This wording is misleading. The data itself can be analysed or interpreted but not extrapolated. One is extrapolating the treatment effect, or the outcome measures, from a reference population to a target population (pediatric). One recommendation is to replace, in this sentence and elsewhere in the document, "extrapolation of data" by "extrapolation of treatment effect" or "extrapolation of (safety/efficacy) outcome".	
ISCT	49	49	1,3	Suggest to caveat	Suggest 'MAY' instead of 'CAN'

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
ISCT	54	54	1,3	Suggest to clarify	Suggest 'Although there are some quantitative strategies mentioned or explained within the guideline, THE GUIDELINE is not meant to be a comprehensive instruction guide.'
EFPIA	60	60	1,4	A verb may be missing after further: "… when necessary to further the scientific understanding of a medicinal product's use in children"	The use of pediatric extrapolation ensures that children only participate in clinical trials when necessary to further investigate/assess the scientific understanding of a medicinal product's use in children.
EFPIA	61	66	1,4	Regulatory authorities per default request a pediatric development and sponsors have to argue for a waiver even when it is obvious that the disease does not exist in a pediatric population and even in case of MoA may be applicable to a pediatric indication this will not imply similarity between the adult and ped condition	When a disease condition is existent in adults and a ped. population this implies a degree of similarity between the reference and target (in this case pediatric) population.
EFPIA	62	64	1,4	This sentence is not fully correct as for new products pediatric development plans need to be proposed by the Applicant early during development. Sentence may be revised as proposed.	When regulatory authorities require pediatric subsets as part of adult-driven drug development, the rationale for doing so can implicitly assume a degree of similarity between the reference and target (in this case pediatric) condition or to fill an unmet need that the new drug may be able to address.
Koop Phyto	66	66	1.4	Ratonale see line above	Please add: In addition, empirical approaches based on real world data should be considered in order to use the scientific and regulatory knowledge of already existing drugs for pediatric extrapolation.
Lundbeck	70	70	1.4	Similarity of the "course of disease" is mentioned here, whereas elsewhere, e.g. in Figure 1, "similarity of disease" is mentioned. These concepts are not exactly the same, should it be "disease" in line 70?	Replace "course of disease" with "disease"
EFPIA	77	82		It is appreciated that the guidance emphasizes selection of designs that address uncertainties, rather than using discrete categories of extrapolation. However, it is difficult to understand what design elements should be considered in what situations. For example, could one always forego an adequate and well controlled study in a target population, so long as there is some low-quality data, supporting similarity in disease/response between target and reference populations?	Add more detail on which specific clinical trial design elements need to be considered and in which situations.
Takeda	78	79	1,4	Other than tolerable level of uncertainty, consideration of the overall benefit risk with the extrapolation to the target pediatric population should not be forgotten.	
ISCT	79	80	1,4	Suggest to clarify	Suggest: 'WHERE TRIALS ARE REQUIRED, options for trial designs will depend on the level of uncertainty that needs to be resolved.'
ISCT	79	79	1,4	Suggest to remove the reference in line 79 to 'Figure 1' as the figure is not explained until section 1.5.	Suggest to remove the reference in line 79 to 'Figure 1' as the figure is not explained until section 1.5.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Agios	79	80	1,4	Within the same trial design there can still be multiple options for the statistical or other methodology that will be employed. Expand sentence to call this out.	Options for trial designs and methodology will depend on the level of uncertainty that needs to be resolved.
GPT	82	82		The figure is appreciated to get an overview. It is, however, as announced, not comprehensive. It lacks principal approaches that are mentioned otherwise in the guideline.	We recommend adding real world evidence data as an option in the pediatric extrapolation plan as a potential study design basis.
German Pharmaceutical Industry Association	82	82	1	The figure is appreciated to get an overview.	
AESGP	82	82	1	The figure is appreciated to get an overview.	
EFPIA	82	83	1,4	Several comments in regards to Fig 1 which would need some reorganization and addition: 1. expand tiing together level of uncertainly and design elements to other parts, eg Tbl. 1 2. take case example, abstract and include so it gets clear, what aspects need to be discussed 3. why are you using a double arrow? What does the colour mean in each graph? How does one go from the concept to plan? Why are similarity of disease and evidence to support similarity both horizontal axes (do they usually correlate with either both of them being in red or both being in green?) 4. An additional suggestion is to illustrate this in a 2-d graph. For instance, if the two main dimensions influencing the extrapolation plan are the similarity and the weight of evidence, those can be shown in a 2d graph (e.g., horizontal is the similarity, and vertical is weight of evidence). Thus, the extrapolation plan continuum would have multiple quadrants. 5. illustrate with a few examples how such graph could be used in submissions or in regulatory decision making. 6. propose disconnecting the third arrow as this is related to the extrapolation plan (which can still allow/account for differences/gaps), rather than the extrapolation concept which is then represented by the 1st 2 arrows in Figure 1. 7. first bidirectional arrow toward the right: Does response to treatment refer to existing drugs or drug classes or only to the investigational drug? 8. it is not clear from the guidance when which design option would be best. Provide more specific guidance on choice of study designs in Section 4.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	82	83	1,4	9. Where is generalization covered in this graph - if there is no additional pediatric data needed to be collected at all in any age group and simulation could be used? We suggest providing alternative examples for the high confidence level in extrapolation for the "Potential Study Designs" image: e.g., omitting additional pediatric efficacy and/or safety studies. Or, if "Exposure matching" means omitting additional pediatric studies, we suggest clarifying. See the general comment, above, on Exposure Matching. 10. Bayesian strategies and "modified" frequentist approaches are talking about the methodology while Figure 1 is more strategic, we do not think those 2 are at the same level than others. 11. Relevance of Figure 1 is unclear e.g is there any ranking of these type of data? e.g Have RWD/RWE the same rank/status as clinical trial data? 12. The term "exposure" used alone in the bottom right arrow is not clear. "Exposure" does not always mean "PK exposure". Exposure could means for example the duration a patient is exposed to a given dose. 13. Consider using alternate colors/shading in Figure 1 to avoid the more common forms of colorblindness	
GPT	86	115		The stepwise approach of an extrapolation concept preceding an extrapolation plan is supported.	
German Pharmaceutical Industry Association	86	115	2	The stepwise approach, extrapolation concept first, followed by an extrapolation plan is strongly supported. The initial focus on existing information should be beneficial for the planning step.	
AESGP	86	115	2	The stepwise approach, extrapolation concept first, followed by an extrapolation plan is strongly supported. The initial focus on existing information should be beneficial for the planning step.	
ISCT	93	95	2	Rationale: The paediatric concept is developed after the review process is completed.	Suggest to move "Once a review of the existing knowledge has been conducted, the data should be synthesized to develop the pediatric extrapolation concept." to end (i.e, after current line 98)
ISCT	96	96	2	Suggest to add additional line at end: 'WHERE FEASIBLE THE EXTRAPOLATION CONCEPT AND PLAN (OR PARTS THEREOF) SHOULD BE VALIDATED PRIOR TO EXECUTION'.	
Lundbeck	103	106	2	This flexibility regarding updating the extrapolation concept and plan is very good from a scientific perspective, especially given that a PIP is often provided prior to finalization of an adult progran and the likelihood is high that more information will become available. However, there is also a risk that this will become a never ending story. How should this be applied in practice? When would one conclude that sufficient information has been gathered to make a conclusion? Also, what should be done if new information becomes available while a pediatric plan or study is being executed?	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
German Pharmaceutical Industry Association	103	106	2	The explicit option to adapt an existing extrapolation plan due to new clinical and scientific data is strongly suported.	
AESGP	103	106	2	The explicit option to adapt an existing extrapolation plan due to new clinical and scientific data is strongly suported.	
Agios	105	106	1,4	This is ambiguous and can lead to misinterpretation. The plan can be modified but should be clear that this is before, for example, a database is locked or unblinded for the final analysis of the pediatric data or extrapolation. Assuming that the guidance is not trying to convey that the final analysis for the purpose of regulatory decision-making does not have to follow a plan/protocol and can change what is outlined in the plan/protocol after the data are known, the wording in this sentence should be qualified to ensure that ICH E9 principles are still followed.	Rather than abandon an existing pediatric extrapolation plan based on a prior concept, the plan itself can be modified (while abiding to ICH E9 principles) to reflect current scientific and clinical understanding.
Koop Phyto	109	109	2.	The extrapolation concept should also consider the drug safety in the reference and target population	Add this point as 4th point at the top of the figure: "Similarity of adverse reactions and undesired events"
Lundbeck	109	109	2	Figure 2 is somewhat hard to follow. Could a clear starting point and end point be added? At the moment it seems to indicate an endless loop (see also previous point), perhaps add a box "conclusion reached"? Aslo, could it be described (in the text, not in the figure) more clearly in what kind of situations additional data would need to be collected prior to defining an extrapolation plan as it seems that in most cases new data would be collected as part of the extrapolation plan.	
German Pharmaceutical Industry Association	109	109	2	The figure gives a nice overview and will be helpful for communication.	
Pharmetheus AB	109	111	2	Figure 2: "Data generated do not completely address knowledge gaps" The word "completely" is too strong and absolute, hence no arrow out of cycle	Suggest changing "completely" to "sufficiently" and add arrow for yes to make inference on pediatric population.
Agios	109	109	1,4	The figure implies that an extrapolation plan (which in some cases may entail a pediatric study) can me modified after the final data are known. This would appear to be a contradiction with ICH E9 principles	Revise figure or add a footnote to clarify that ICH E9 principles should be followed.
AESGP	109	109	2	The figure gives a nice overview and will be helpful for communication.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	109	111	2	Several comments in regards to Fig 2 which would need some reorganization and addition: 1. suggest replacing "synthesize available data" in this sentence with "synthesize evidence" or "integrate relevant information". In most of the document, "data" refers to clinical data. However, information supporting disease similarity may rely on mechanistic models, published results, and some qualitative evaluations. Thus, in this section of the graphic, we suggest being more inclusive of the source of information beyond clinical data, to avoid confusion. 2. Figure 2 (iterations as new information becomes available) suggests that going from extrapolation concept to extrapolation plan is an iterative process that is updated when new information becomes available. However, how this iteration plays out is not mentioned in individual sections and it would be helpful to illustrate that with examples. 3. suggest adding a note for when the gaps were solved, and extrapolation is completed. Please include an arrow to a box indicating "data generated completely addresses knowledge gaps". 4. Why is the left box ("need for additional data collection") not part of the Extrapolation plan? Why should the concept be changed and not the plan in case of "data generated do not completely address knowledge gaps"? 5. include additional arrows out of the "Execution of Extrapolation Plan." For example, additional outcomes (arrow or arrows to pediatric authorization) may include: - omitting an additional pediatric study (no additional data needed), - conducting a more efficient pediatric bridging study, or - conducting a fully powered well controlled study. We also suggest the "Extrapolation Concept" include that any data and knowledge gathered in the successful execution of extrapolation feedback would inform future compounds in the same disease.	
EFPIA	112	115	2	In certain circumstances it may not be possible to fully confirm assumptions (one can still assess quantitatively but no firm conclusions can be drawn). In the case of rare/orphan diseases or disease subtypes, patient data may be too sparse to allow analyses with sufficient precision.	Suggest to say 'evaluation' of assumptions rather than 'confirmation', stressing also the impact of quality and quantity of data generated in the extrapolation plan. The extrapolation concept can be continuously updated with emerging data from many sources (e.g. data other than clinical trials like real world evidence)
Lundbeck	114	115	2	What is the purpose of this review? Who are the stakeholders. Is it meant to be a "lessons learned" overview for internal use, for regulators, or should this be shared with other companies? Could you please elaborate?	
EFPIA	117	177	3	Entire section: I am missing advice/guidance on when the concept should be developed. Usually, the PDP is initiated after start of the adult program, but a lot of the data collection advice may then be lost if relevant e.g., biomarker data (which may account for the differences in the populations) have not been collected sufficiently in the reference population.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
ProPharma Group	118	120	3	It would be helpful to understand whether the expectation is that the Pediatric Extrapolation Concept should be submitted to Regulatory Agencies for review as part of the paediatric plan. Is it anticipated that this document would be appended to a PIP/iPSP/PPSR? Line 736 talks about the modelling and simulation plan should be generated for internal documentation purposes but is also suitable for interaction with regulators. Is the same true for the extrapolation concept and plan?	n/a
EFPIA	119	119	3	The word "influence" seems to indicate a causal relationship which is sometimes difficult to establish. Understanding the factors that are associated with the similarity of disease may be more appropriate.	Development of a pediatric extrapolation concept requires an understanding of the factors that influence are associated with the similarity of disease, the pharmacology of the drug and the response to therapy as well as the safety of use in all the relevant populations.
German Pharmaceutical Industry Association	122	128	3	The differentiated approach is highly appreciated as it takes into account the biological diversity of a pediatric population.	
AESGP	122	128	3	The differentiated approach is highly appreciated as it takes into account the biological diversity of a pediatric population.	
GPT	128	128		The more nuanced approach is highly appreciated as it reflects more adequately the biological broadness in particular in the concerned vulnerable population.	
German Pharmaceutical Industry Association	130	141	3	The "open" interpretation of disease similarity is strongly supported.	
AESGP	130	141	3	The "open" interpretation of disease similarity is strongly supported.	
EFPIA	138	141	3,1	"anatomic congestive heart failure" may be difficult to understand. Could it be specified to, e.g., "congestive heart failure (CHF) due to unrepaired or palliated congenital heart disease (CHD)"?	For example, anatomic congestive heart failure (CHF) in children due to unrepaired or palliated congenital heart disease (CHD) is not similar to adult heart failure,…
EFPIA	143	148	3,1	Evaluation of disease similarity is not a one-time exercise and as knowledge is gained, the information can be incorporated into the evaluation of disease similarity in the pediatric extrapolation concept. In most cases sponsors must have an agreed PSP and/or PIP with the Regulatory Authorities prior to submitting a marketing application. The guideline is silent on how often or at what frequency sponsors should assess disease similarity and how changes to the pediatric extrapolation concept could impact the pediatric extrapolation plan once the plan is in the “execution phase”.	Add recommendations or provide guidance for sponsors around the frequency of assessing disease similarity. Alternatively, state that the frequency of assessing disease similarity may be disease specific and dependent on the available data.
Lundbeck	146	148	3.1	How to handle this from a practical point of view? Would regulators need a yearly update or an update on request etc.? This is probably something to discuss when submitting the PIP, but it would be nice with some guidance.	
EFPIA	150	150	3.1.	There is a section 3.1.1 but no section 3.1.2	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	154	164	3,1	For oncology, it is important to mention genotypic expression or tumor specific mutations as it is a very important factor for similarity/unsimilarity.	"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, genotypic expression, tumor-specific mutations, or other age-related differences."
Takeda	155	156	3.1.1	How important is etiology for the evaluation of the similarity? Etiology seems less important than other factors since there are many diseases which have different etiology between adult and pediatric populations.	
Takeda	156	157	3.1.1	Physiologic differences and/or similarities (i.e. maturation of renal function) between reference and target population should also be considered.	
EFPIA	160	160	3.1.1	insert "age-related ontogeny".	
EFPIA	163	164	3.1.1	Untreated disease is often not very well studied if treatments are available.	Suggest to amend the wording to read: "Similarities in the outcome of untreated disease should also be evaluated, if possible".
GPT	166	187		These questions seem to describe the disease characterising factors by limiting them to pathophysiological or pathobiochemical factors.	We propose to extend the scope of questions to psychosocial disease determinators.
German Pharmaceutical Industry Association	166	187	3	These questions seem to reduce the disease characterising factors to pathophysiological or pathobiochemical factors.	Extend the scope of questions to psychosocial disease characteristics.
AESGP	166	187	3	These questions seem to reduce the disease characterising factors to pathophysiological or pathobiochemical factors.	Extend the scope of questions to psychosocial disease characteristics.
GPT	170	170		We support the validity of the question. However, this limitation does not reflect the fact that competent authorities impose indications that are not covered by standard diagnostic criteria.	We recommend inclusion of indications not covered by standard diagnostic criteria.
EFPIA	172	177	3.1.1	Suggest being specific that manifestations could refer to severity/intensity of the disease and symptoms.	When evaluating similarities and differences between reference and target populations, the following should be considered (note that manifestations could refer to severity/intensity of the disease and symptoms):
Takeda	205	207	3.1.1	It seems difficult to understand the sentence "What effect have these treatments (e.g., timing of treatment relative to onset of disease and age of the patient, frequency of treatment, length of treatment) had on the course of the disease in the reference and target populations?"	Are there similarities or differences in these treatments (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?
EFPIA	205	205	3.1.1	We assume this is supposed to be a sub-point of the immediately preceding point?	Indent the point to reflect that it goes along with the immediately preceding point or include it as part of the immediately preceding point. If it's not meant to go with the immediately preceding point, make clear what 'these treatments' is referring to.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
German Pharmaceutical Industry Association	209	215	3	Again, the "open" interpretation of disease similarity is strongly supported.	
AESGP	209	215	3	Again, the "open" interpretation of disease similarity is strongly supported.	
EFPIA	209	215	3.1.1	Clarification added: Pharmacometric models of disease progression dynamics leveraging available adult and pediatric data can enable quantitative assessment of similarity in disease trajectory and identify intrinsic or extrinsic factors that influence disease dynamics to inform the pediatric extrapolation plan.	Addition: Pharmacometric models of disease progression dynamics leveraging available adult and pediatric data can enable quantitative assessment of similarity in disease trajectory and identify intrinsic or extrinsic factors that influence disease dynamics to inform the pediatric extrapolation plan.
GPT	217	234		For substances which are biologically defined (e.g,. biologicals, herbals) or topically applied products such as vaccines, PK data cannot be generated and are sometimes not even useful, because e.g. they do not represent the entire active ingredient. Nonetheless, e.g. herbal medicinal products have been used successfully for many decades and are often well studied scientifically.	The PK/PD-based approach is only one option among several. If not available or not possible/feasible/reasonable, other tools such as empirical approaches need to be used, e.g. for locally applied medicines with local effects, for herbal medicinal products or for vaccines.
EFPIA	217	235	3,2	In section 3.2 there is no mentioning of similarity of exposure-response beyond in the last sentence. I recommend that the section should include a discussion on similarity of exposure-response. This is mentioned in section 3.3, but I believe it belongs here and it should be discussed in more detail as currently done in section 3.3. Mainly two points need to be discussed: (1) Similarity of exposure-response is closely linked with baseline risk factors. If there are no risk factors that need to be considered, then similar exposure should lead to similar efficacy. But if there are additional risk factors (like baseline severity) to be considered, then similar exposure will only lead to similar efficacy conditional that the risk factors are the same. This type of discussion is almost completely missing in the guidance (only briefly mentioned in lines 842-847 or 868-876), whilst it is absolutely essential and should be the core of the guidance. (2) one often can't even establish an exposure-response relationship in adults, let alone in children. Hence, similarity of exposure-response can't be demonstrated by comparing two exposure-response relationships. When there is only an exposure-response relationship established in adults, but not in kids, one may still be able to demonstrate similarity of response for the exposure-level that is efficacious in adults. If (like in many oncology indications) there is only data on the approved dose (or just sparse data for other doses) then an exposure-response relationship may not be available even for adults. What can we do to demonstrate some level of similarity in such a case?	
EFPIA	217	234	3,2	The considerations around ADME properties are definitely important for pediatric development, but are more about determining what dose is appropriate for different age groups rather than disease similarity (as opposed to MOA differences, which could be relevant for both). It seems like there a couple of concepts mixed up in this section (how to select the right dose vs how to assess disease similarity), which could be clarified. Propose to change "drug pharmacology similarity" to "PK/PD Similarity"	It seems like there a couple of concepts mixed up in this section (how to select the right dose vs how to assess disease similarity), which could be clarified.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	217	237	3.2-3.3	For clarity, more clarification of the different categories is requested. Consideration might also be given to moving some elements from Section 3.3 to Section 3.2 e.g. mechanism of action.	Integrate Sections 3.2 and 3.3
Takeda	236	245	3,3	Should also consider surrogate endpoints that are non-traditional endpoints used in the reference population.	
EFGCP	236	284	3,3	Consideration of Quality of Life in treatment response/response end points	Consideration of whether the evidence suggests that the QoL implications of treatment are different in the paediatric population (e.g., greater impact on those factors important to the paediatric population (school/friends) in comparison to the adult population).
EFPIA	236	284	3,3	It is not clear whether the term "similar response to treatment" refers to the treatment under investigation, or whether it refers to all treatments in general. This should be clarified, and these two aspects should be dealt with separately. Much of the content of the section (for example when discussing endpoints) seems to refer to treatment in general. Similarity of the response to the treatment under investigation is somewhat limited here to similarity of exposure-response and fits better in section 2. In this section (and elsewhere in the guidance) the terms "PK/PD relationship" and "exposure-response relationship" are being used. It is not really clear whether these terms are to be understood as synonyms (in that case only one of the two terms should be used throughout the document) or if they describe two (maybe slightly) different concepts (in that case a definition of the two terms would help).	
GPT	237	245		Emphasizing the continuum in disease response is appreciated.	We suggest extending the scope to factors that include patients' expectation while managing the care of the disease.
German Pharmaceutical Industry Association	237	245	3	Viewing disease response as a continuum is supported.	
AESGP	237	245	3	Viewing disease response as a continuum is supported.	
EFPIA	237	237	3,3	The second comma appears to be a typo; suggest removing (or confirm it is intended to be in the sentence).	
Koop Phyto	238	238	3.3	As with similarity of disease, the similarities, and differences in response to treatment between a reference and target population should be understood as a continuum	At least a reference to this should be included, or better, an assessment should be requested whether a continuous process is to be expected.
Pharmetheus AB	241	243	3,3	"Similarly, data generated in other indications for the drug can serve as a relevant source of knowledge when assessing the similarity or difference of response to treatment."	Assuming response to treatment similar across indications seems like a very strong assumption, thus we recommend some phrasing of caution and wording concerning how such assumption is validated or under what circumstances it is reasonable.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Takeda	241	243	3,3	Recommend to expand on the sentence "Similarly, data generated in other indications for the drug can serve as a relevant source of knowledge when assessing the similarity or difference of response to treatment." by adding text regarding the mechanism of action.	Similarly, as far as the same mechanism of action, data generated in other indications for the drug can serve as a relevant source of knowledge when assessing the similarity or difference of response to treatment.
EFPIA	243	245	3,3	This statement seems to imply that the assumption of similar exposure-response is required. This should be stated more clearly, as it is important for the paediatric dose selection (which is usually a dose that achieves the same exposure as observed in adults when treated with the registered adult dose). Also, the guidance should discuss in a bit more detail how to handle different cases. In many indications a proper dose finding study cannot even be done in adults, so that our knowledge about the adult exposure-response is limited (case 1). In other indications, we may have an adult exposure-response, but we can only collect limited data in children (by using just one paediatric dose regimen that matches the exposure observed under the registered adult dose regimen [case 2]). The case where we can assess similarity of exposure-response with adequately characterized E-R relationships in both populations (case 3) is probably rare.	
EFPIA	247	247	3.3.1	No secction 3.3.2?	
EFPIA	252	253	3,3	The current section implies that PK/PD must be done in a single step, however we often establish a PK model in one step and then a PD model in a second rather than doing a single joint PK/PD analysis. The text suggests that only a single joint PK/PD analysis is intended, but that is not easy to do in many instances. This is why a 2-step procedure is often used.	
Lundbeck	257	258	3.3.1	Minor: this example is very obvious. Perhaps use a little less obvious example to illustrate how expected response can depend on age?	
EFPIA	257	258	3.3.1	The text states that "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." This is a very extreme scenario/example and, as such, it is suggested that a reference be included or the sentence be deleted.	It is suggested that a reference be included or the sentence be deleted

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	258	261	3,3	Indeed, if such factors exist and have a direct impact on response, similarity of exposure-response and response to treatment should be evaluated within the factor-levels of these "baseline risk factors". This point should be discussed and elaborated in more detail, either here or elsewhere in the guidance. With such additional factors present, similarity of response should apply for patients who do suffer from a comorbid disease, and it should apply to patients who do not suffer from a comorbid disease. Adult and paediatric patient populations may differ with respect to the baseline distribution of this risk factor (i.e., many more adults suffer from the comorbid disease as compared to children), but extrapolation should still be possible within the subgroups (with / without comorbid disease). This should include exposure-response plus these additional risk-factors. There are more statements earlier in the document also (lines93, 118) that state factors that influence effects of treatment should be identified, but I did not see anywhere in the document what should then be done with that. How does it impact the discussion on extrapolation etc?	
Koop Phyto	261	261	3.3.1	The PK/PD based approach is not suitable e.g. for locally applied medicines with local effects, vaccines or herbal medicinal products. Other tools such as empirical approaches need to be used and accepted. As a retrospective approach, the systematic collection of the existing experience from previous therapeutic use in children in a scientific manner can provide valid information on the safety and the therapeutic usefulness of such products.	Please add: If meaningful PK/PD data are not available (e.g. for multicomponent herbal preparations, local acting drugs or vaccines) other data sources such as real world data or data from longterm medical experience documented in surveys, case reports or preclinical studies can be used. Given the wide-spread off-label-use in paediatric real life, research including data collection and evaluation should listen to these users and their experience, empirical studies based on real-life data are a considerable option for such cases.
GPT	264	276		The list is not complete with respect to indications that are not defined by biomarkers or standardized questionnaires.	We recommend adding new endpoints or questionnaires for the indications concerned.
EFPIA	264	275	3,3	Please provide additional guidance explaining why these questions are crucial to consider when evaluating the similarity of response to the endpoint(s) and discuss the consequences that might result from different possible answers. Currently, the document only lists the questions without providing any guidance. We also suggest adding a note that these are not the only type of questions that Sponsors should evaluate but rather a recommendation.	
EFPIA	269	270	3.3.1	We do not see the difference versus 1st bullet point just above.	
EFPIA	274	275	3.3.1	"to a biomarker endpoint in the target population" Is this sufficient to warrant extrapolation, we are wondering if we should extend to a bit more the infomation needed?	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	277	279	3,3	Please include language about the impact of effect modification of some characteristics (say, by age, and weight) on the evaluation of similarity. That is, for example, if in the reference population, the magnitude or direction of the treatment effect (contrast between test and control) varies by age or weight, then these factors interact with treatment on response in the reference population. What would be the impact of such interaction on the extrapolation plan to the target pediatric population?	
Takeda	279	281	3.3.1	Not sure about the sentence "For many pediatric drug development programs, the primary endpoint(s) in the target pediatric population is/are different from that in the reference population." Please expand and/or clarify.	
EFPIA	279	283		Based on prior experience (across the industry), there is significant risk to the success of a pediatric study if the primary endpoint is different for pediatric vs adult studies. Some discussion (or at least acknowledgment) of this would be helpful here, though we fully acknowledge that it is not always possible or appropriate to use the same endpoints.	Some discussion (or at least acknowledgment) of this would be helpful here, though we fully acknowledge that it is not always possible or appropriate to use the same endpoints.
Lundbeck	281	282	3.3.1	A comparison of a common component or combination of common components would result in a comparison of potentially unvalidated endpoints, would that be acceptable? Is it required that some kind of validation be done? Likely not all aspects of the endpoint would be covered if focus is on certain components only?	
EFPIA	281	283		This sentence is not clear in what is being suggested. Is this suggesting evaluating the correlation between different endpoints based on data measured in the reference population, considering the similarity/scaling to the endpoint to be used in the target population? The guideline mentions that many of the primary endpoints in the target pediatric population are different from that in the reference population. Additionally, the guideline states that a comparison of one or more components of the primary endpoints and/or secondary endpoints can be used to understand the relationship between the different endpoints. However, the guideline is silent on approaches sponsors can utilize to compare different endpoints in the reference and target populations and how the information should be interpreted.	Suggest adding more clarity to this recommendation. Include guidance/strategies for evaluation and interpretation that sponsors can consider when the endpoints in the reference and target populations differ.nclude guidance/strategies for evaluation and interpretation that sponsors can consider when the endpoints in the reference and target populations differ.
GPT	285	299		Data that lead to market access of medicinal products or which form a regulatory basis for registration e.g., European herbal monographs established by the HMPC, or equivalents in China, Australia or Canada, are not mentioned.	We recommend including such data as well, also in the table under "Other sources". To our understanding real world evidence can also consist e.g. in NIS, registries, or prescription data.
German Pharmaceutical Industry Association	285	299	3	Data that lead to market access of medicinal products, e.g., European herbal monographs established by the HMPC, are not considered.	We recommend to include such data as well.
AESGP	285	299	3	Data that lead to market access of medicinal products, e.g., European herbal monographs established by the HMPC, are not considered.	We recommend to include such data as well.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
ISCT	289	289	3,4	Not all data will be of the required applicability or quality to be used	Suggest that "All available data should be EVALUATED FOR SUITABILITY FOR CONSIDERATION IN establishING the extrapolation concept and formulate the extrapolation plan."
EFPIA	290	291	3,4	Clarification added: Such information may also include data from ongoing adult/pediatric development programs, or relevant data from terminated programs, including competitor with same MoA.	Such information may also include data from ongoing adult/pediatric development programs, or relevant data from terminated programs, including competitor with same mechanism of action.
GPT	297	298		We appreciate and highly support that real world data are included as this reflects the actual use of many medicinal products.	
Koop Phyto	297	299	3.4	These are also important scientific sources of RWD	Table 1: Please add to Real World Data: Periodic Safety Update Reports, regulatory or scientific monographs
German Pharmaceutical Industry Association	297	298	3	We appreciate and highly support that real world data are included.	
AESGP	297	298	3	We appreciate and highly support that real world data are included.	
EFPIA	297	299	3,4	Although the table lists the sources, not the motivation for data collection, and IIT & off-label data should be captured in various types of data sources, IIT/off-label data and the results from published papers are not explicitly mentioned within the guidance. This is a crucial data point, and it would be beneficial to explicitly state that these are acceptable data sources. Clinical Data; second row: PK, PK/PD, E-R, and clinical data in other related conditions for a drug or drugs in the same class. This is critical and repeated on line 298 of this section, what does the word related mean? Does that mean within a condition such as immune diseases e.g., psoriasis, JIA, SLE, MS. What about HS, Lupus nephritis, urticaria? How much information can we leverage – is it only PK/PD data? Can the information be used for dose, safety, efficacy? Line 555 in 4.1 is very important and supports this as well.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	297	297	3,4	Several comments regarding Table 1: 1. Suggestion revising the title of the Tabl, as the table is not only about sources for "similarity of disease", but also similarity of disease, pharmacology and response to therapy: "Examples of Sources and Types of Data to Evaluate for Similarity of Disease and Response to Treatment Between Reference and Target Population" 2. Suggestion adding a row for "data from tumor banks/assessments of relevant biomarkers." 3. Clinical: Consider adding explicit mention of modeling and simulation, similar the "in silico" methods for nonclinical data, lines 309/310 (to align with non-clinical row); Row: "Clinical Data" may benefit from combining the sub-rows under "Types of Data" as it is clearer to list the same disease condition for both the same class and a different class: suggest combining the first and third sub-rows under "Types of Data" to one row to simplify the Table. 4. Other sources: "published" does not qualify a model, however the formulation seems to exclude other models (e.g. qualified and documented internal models). Require instead that a model is qualified for the context of use and documented. 5. Why are only meta-analyses or reviews as literature source are considered as a source and not other published papers. It is suggested that the text allows for published literature to be considered also: "Relevant published literature, systematic reviews or meta-analyses including those that can be used to evaluate suitable biomarkers."	
ProPharma Group	300	305	3,4	From previous experience with the EU Agencies when proposing an extrapolation approach, presentation of clinical data available to date (phase 2 in adults) to demonstrate the similarity between adults and children has been met with comments that without phase 3 adult data being available, it is difficult to predict to what extent these data will be useful in predicting extrapolation of efficacy to children and therefore the ability to compare and extrapolate has been questioned. It would be helpful if the guideline could clarify in general how much clinical data is expected to be able to agree an extrapolation approach. With regards to agreeing paediatric plans in the US and EU with Regulatory Agencies, these plans would be considered late if phase 3 data were awaited before agreeing the approach. Therefore there appears to be some disconnect between data required and timings per guidance.	n/a
GPT	300	305		Combinations products are not specifically mentioned.	For combinations we suggest accepting available data for single components as well as e.g., the evidence for the combination when evaluating clinical data.
German Pharmaceutical Industry Association	300	305	3	Combination products are not specifically mentioned.	We recommend for combinations to accept data for single components as well as data for the combination when evaluating clinical data.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
AESGP	300	305	3	Combination products are not specifically mentioned.	We recommend for combinations to accept data for single components as well as data for the combination when evaluating clinical data.
EFPIA	301	302	3,4	"Clinical data (e.g., from controlled trials, prospective observational studies, PK, PK/PD and/or biomarker studies) in populations with the same condition or related conditions should be evaluated to understand similarities and differences between the reference and target populations." Tumor banks with assessment of relevant biomarkers should be included in the text.	We propose the following addition: "Clinical data (e.g., from controlled trials, prospective observational studies, PK, PK/PD and/or biomarker studies, tumor banks with assessment of relevant biomarkers) in populations with the same condition or related conditions should be evaluated to understand similarities and differences between the reference and target populations."
EFPIA	303	303	3,4	More clarifications/examples should be included in order to explain what is intended by the term 'condition' (as opposed to 'disease').	
Agios	304	305	3,4	Patient-reported outcomes data can provide unique information that speaks to the impact of treatment from the patient's perspective and can assist with the interpretation and contextualization of observed effects from more traditional safety and efficacy measures.	All available data, which could include patient-reported outcomes data, for the drug/drug class should be evaluated including ongoing and completed studies, published or unpublished, whether results are positive or negative.
EFPIA	308	309		While the emphasis on using all available data is appreciated, inclusion of nonclinical in vivo, in vitro, and in silico models may need to be vetted to avoid misleading results. Mechanistic, and semi-mechanistic models employing in vitro and in silico data are only as good as the input data. Some guidance on how to assess the quality of the data would be valuable to avoid unnecessary delay resulting from building time-consuming models that may be of limited value or even mid-leading.	
German Pharmaceutical Industry Association	316	322	3	Discussion of RWD with regulatory authorities is supported.	
AESGP	316	322	3	Discussion of RWD with regulatory authorities is supported.	
Koop Phyto	317	318	3.4	The scientific importance of RWD should be accepted and not only be considered.	Please change the sentence to: Therefore RWD can be used to support pediatric extrapolation.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	318	337	3,4	While there may be benefit in discussing the RWD strategy with regulatory authorities, this may also be appropriate for other aspects of the pediatric strategy such as the full extrapolation plan, (as mentioned later in the guidance). As such, it might be more appropriate more sense to include a more general statement in line 337.	318: The extent to which RWD can be used to support pediatric extrapolation, both the pediatric extrapolation concept and plan, is evolving. Thus, consideration should be given to Therefore, the adequacy, and relevance of such data, and extent to which RWD can be used to support pediatric extrapolation should be discussed with regulatory authorities. 337: A critical and multidisciplinary assessment of all the data should be conducted to justify the use of the evidence to support the extrapolation concept. Consideration should be given to discussing the pediatric strategy, including sources and types of supporting data, with regulatory authorities.
Koop Phyto	320	320	3.4	These are also important scientific sources of RWD	Please change the sentence to: […] but not limited to electronic health records, claims databases, registries, regulatory and scientific monographs and periodic safety update reports, can be considered
EFPIA	339	339	3,5	The Guideline uses "a priori" information for efficacy evaluation and extrapolation. We suggest adding similar language to Section 3.5.1. Extrapolation of Safety for methods that use a priori for safety evaluation	It is helpful and relevant to know how much of the data that has been generated in the reference population could be used to improve the interpretability of the safety data in the target population.
Agios	340	341	3,5	Please also refer to ICH E9 which has dedicated considerations as they pertain to analysis of safety data	Basic considerations for the development of an overall safety data collection and adverse event reporting plan are discussed in other guidances (ICH E2, ICH E6, ICH E9, ICH E11, ICH E11(R1)).
EFPIA	357	357		"The source and amount of safety data to support the extrapolation of safety data to a target population", It is suggested to remove "data" in the second occurence	The source and amount of safety data to support the extrapolation of safety data to a target population.
EFPIA	358	362	3.5.1	Several comments regarding including more examples on safety extrapolation: 1. When discussing extrapolation of safety, the guidance mentions that data can be leveraged in reference populations who have been treated with different dosing regimens and/or different diseases/indications. Generally, the extent of extrapolation is partly based on similarity of disease and response to treatment, therefore, it is unclear what data from different dosing regimens and/or diseases can be leveraged from reference populations who have different diseases. 2. Clarify under which circumstances one can extrapolate safety as disease will be important covariate	Recommend to include examples of what data can be leveraged in reference populations with different dosing regimens and/or diseases to support extrapolation of safety and strategies on interpretation.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	361	362	3.5.1	The text currently limits inclusion to adolescents. Consider adding additional text that could facilitate an understanding when even broader inclusion to children could be considered appropriate.	
EFPIA	361	363	3.5.1	"Enrollment of adolescents in/or concurrent with the adult trials may allow for earlier evaluation of safety for the adolescent population." We would suggest changing the sentence.	Enrollment of adolescents in,/or concurrent with, the adult trials may allow for earlier evaluation of safety for the adolescent population
EFGCP	366	415	3,5	Consideration of the delivery mechanism for the treatment in question, e.g., is this a tablet that will have to be crushed for the paediatric population because there is not an appropriate formulation? Will this give rise to any further safety concerns?	Add the aspect of crushing tablets as a consideration in the safety section to make it an explicit consideration for the paediatric population
EFPIA	366	392	3,5	Please provide additional guidance explaining why these questions are crucial when considering extrapolation of safety, HOW the answers to these questions may impact the development plan, also discuss the consequences that might result from different possible answers, and the additional information that may be required. Currently, the document only lists the questions without providing any guidance. We also suggest adding a note that these are not the only type of questions that Sponsors should evaluate but rather a recommendation.	
ISCT	375	375	3.5.1	Suggest to add ON-TARGET OFF-TISSUE EFFECTS as an additional consideration because sometimes the molecular target may be correct but adverse events may occur if the correct tissue is not targeted (e.g., in gene therapy)	
EFPIA	375	375	3,5	insert "are the known safety effects mechanism related?"	
EFPIA	375	375	3,5	It is acknowledged that the factors mentioned on page 21 of the guideline are important considerations. However, depending on the issue it is also relevant to consider whether there are other drugs with a similar mechanism of action that could provide information on the safety profile to be expected. This could be safety related to primary pharmacology or off-target effects.	Consider adding the following question: What is known about the pediatric safety profile of other drugs with similar on- and/ or off-target activities at pharmacologically contextualized exposures (e.g., at similar multiples of unbound exposures relative to target or off-target potency)?
EFPIA	380	380	3.5.1	"treatment effect size" is usually understood as the efficacy of the treatment compared to placebo or a reference treatment. Not mix it with treatment duration, this could be a separate bullet and allude to the magnitude of the effect size. Separate bullet and allude to the magnitude of the effect size	
EFPIA	383	385	3.5.1	The question about the expected drug exposure should also be added to the section 3.3.1.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	394	395	3.5.1	Could an example of a drug approved using any of the extrapolation methods mentioned in the above bullets be provided as an example. Or an example of extrapolation of safety from adolecents to other age groups such as children?	
EFPIA	395	397	3.5.1	Please explain these "circumstances" in more detail, beyond there having to be confidence.	
Koop Phyto	397	399	3.5.1		Many safety data require much higher case numbers than efficacy evidence and are therefore also primarily unlikely to be fully clarified in pre-approval studies. Therefore, collecting safety data for newly developed substances before and after approval in children may be justified.
EFPIA	399	399	3.5.&	Can examples be added here? we are always asked to collect safety data in pediatric trials.	
EFPIA	401	420	3.5.1	Though the previous section concludes with "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 pre-authorization program" this seems to be doubted with "additional safety considerations" which is also vague in the examples and it may be helpful to learn about undetected safety issues in the ped. Indications so far approved over the last 10 years; are there examples in support of prolonged trials and larger sample sizes only for safety? Sample size is usually defined based on efficacy-what is meant by the sentence "the use of arbitrary sample sizes without appropriate scientific justification is discouraged". The examples seem also not evidence based or if so it would be good to learn about the evidence for "narrow ther index drugs" or when the drug is a new MoA? If safety extrapolation is allowed, there should be a fair discussion between regulator and sponsor but this section opens the door to non-harmonized approaches.	
EFPIA	405	406	3.5.2	Please elaborate further on this example. A reader who is not familiar with this specific situation may not understand why and how "the effect of corticosteroids on reduction in growth velocity" is a good example for the need to collect additional safety data.	
Takeda	406	407	3.5.2	Does this sentence indicate the longer-term safety in target pediatric populations data will not to be required as clinical data package if there are no remaining gaps and/or age-specific long-term safety concerns in the target population?	
EFPIA	407	407	3.5.2	For very young pediatric population, due to ontogeny changes, the safety profile may change. need to add a statement in.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	409	416	3.5.2	This section addresses special considerations for the collection of pediatric safety data. As mentioned earlier in the document (3.2), ADME properties for a drug can differ between the adult and pediatric population. Please explain these "special consideration" for the four cases in a bit more detail.	Add the following bullet: • When the ADME characteristics for a drug differ substantially in the adult and pediatric population
Lundbeck	411	411	3.5.2	This example is often brought up in the discussion of extrapolation of efficacy as well. It would be beneficial with clarifications as to whether these special considerations also apply to extrapolation of efficacy	
EFPIA	416	416	3.5.2	Consider adding the following example in the special considerations : When co-morbidities in the paediatric population are remarkably different to those in the reference population.	Add: When co-morbidities in the paediatric population are remarkably different to those in the reference population.
Lundbeck	417	420	3.5.2	Throughout the text it is sometimes hard to distinguish what should be part of the extrapolation concept and what should be part of the extrapolation plan. This paragraph discusses study design, does it not belong in a section concerning the extrapolation plan? It is currently part of a section on the extrapolation concept.	
EFPIA	417	420	3.5.2	Please elaborate how study designs might depend on the gaps, maybe with an example. What is an arbitrary sample size? And what is an appropriate scientific justification of sample size. Is a sample size that is based on incidence and prevalence rates (and hence the ability to recruit) arbitrary?	
EFPIA	418	420	3.5.2	The guidance states that the use of arbitrary sample sizes without scientific justification should be avoided, but occasionally regulators might propose a different sample size from Sponsors during pediatric planning, specifically to increase exposures for safety, and in those cases it does seem to be an arbitrary number that lacks scientific justification. Given that in many peds programs studies are not directly sized to address specific safety questions, how does this guidance help Sponsors and Regulators align on sample sizes for purposes of safety? Can anything be added on this subject?	Clarification is requested as to where the standard requirements of ICH E1 (Population Exposure: The Extent of Population Exposure to Assess Clinical Safety) for the size of the pediatrics safety database still apply and, if not, if guidance on a target size of pediatric safety database, based on (dis)similarity between adults and pediatrics can be provided.
EFPIA	420	420	3.5.2	Insert: for some indications, due to recruitment challenges, the sample size may be limited without considering statistical power.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
German Pharmaceutical Industry Association	422	494	3	The idea of a Pediatric Extrapolation Concept is strongly supported.	
AESGP	422	494	3	The idea of a Pediatric Extrapolation Concept is strongly supported.	
EFPIA	422	494	3,6	This section should be a key section of the extrapolation concept bringing all of the information together in a coherent manner, unfortunately, the current text is difficult to understand and even more difficult to see how this could be applied in practice. The first paragraph (lines 423-427) outlines a very agreeable objective, but the rest of the section fails to meet this objective. The 'Assessing Similarity of Disease between pJIA and RA' graph on slides 12 and 13 of the case study is a great real-world example of how to take the information and display it. What we would like to see in the guidance is a description on what are the steps you need to take to get to the graph and some guidance on how discuss similarity of the areas shown (e.g., manifestations of disease, measurements used, subtypes, other factors). From there, there should be a guidance on how one is supposed to make an assessment of what the assumptions are, where are we on the continuum of similarity, and whether this is purely an assessment based on guess or whether this is based on data and what data, etc. Essentially, what does this graph mean and how do we get there. What could be discussed is (by area) how to review evidence, sources of information, how to assess the similarity on the continuum, and how to decide whether there are gaps and how to close these gaps. The presentation of this could be the graphic of the case study (slide 12). "By area" refers to the four different arrows in the graph.	
EFPIA	422	422	3,6	For some diseases, the pediatric extrapolation concept has already been established. For example, partial onset seizures, antibacterials, and antivirals. It would be duplicative and inefficient for each pediatric program to redo the evidence synthesis in such cases	We suggest including that in some diseases, pediatric extrapolation may already be well-established.
EFPIA	433	442	3,6	It is good to have a list of questions that need to be addressed. It would be even better if there would be guidance HOW the answers to these questions impact the development plan, and the additional information that may be required. At least please provide some examples.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Takeda	439	439	3,6	The use of consistent vs similar should be reconciled throughout the document.	
EFPIA	441	442	3,6	This sentence is very difficult to understand, especially the second part: “how do these differences affect assessment of similarity”. What do you mean by “difference” here? Do you mean “inconsistencies”?	
EFPIA	444	447	3,6	Please provide examples on HOW these answers will inform what additional information is recommended.	
EFPIA	444	476	3,6	This section focuses on integrating evidence around the endpoint without much discussion about heterogeneity assessments of the populations that are essential in any evidence integration. Please include guidance on evaluation of heterogeneity of population and findings in the reference population.	
ISCT	446	447	3,6	Suggest adding line space before sub-heading 'Methodologies that can be used to integrate evidence'	
EFPIA	446	447	3,6	Blank line missing.	Add blank line before "Methodologies that can be used to integrate evidence"
EFPIA	449	450	3,6	It is stated that "Use of mechanistic and/or empirical approaches in the synthesis of data should be considered". Since a sound quantitative sysnthesis relays on statistical methods this should be explicitly mentioned.	Rephrase sentence: Use of statistical, mechanistic and/or empirical approaches in the synthesis of data should be considered.
EFPIA	450	450	3,6	The expression "systems biology/pharmacology data" might be ambiguous. Is this about modeling / QSP? Please clarify.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	450	452	3,6	Sentence implies that system biological or pharmacological data should be considered only if population-level data is available. Because all avilable date should be considered sentence should be rephrased to guide for more flexibility here.	Provide list with examples of data to be considered: * Population-level data on efficacy and/or safety * Patient data on efficacy and/or safety * Systhem biological/pharmacological data * …
EFPIA	451	452	3,6	It should be clarified that the parameters in parentheses are examples rather than a comprehensive list. The text should be amended accordingly.	Inclusion of systems biology/pharmacology data from the reference population(s) should be considered when population-level data (e.g. epidemiological, diagnosis and non interventional study data) are valeyable.
EFPIA	452	454	3,6	It is welcome that the importance of quantitative synthesis of existing data is stressed. Since Bayesian methods integrate evidence naturally, the should be listed as an option in the paragraph. Please revise the sentence to include safety. "Meta-analytic techniques for synthesizing efficacy and safety data in the reference population(s) should also be considered."	Expand sentence: "Meta-analytic and/or Bayesian techniques for synthesizing efficacy and safety data in the reference population(s) should also be considered".
EFPIA	452	453	3,6	Clarify whether meta-analytic techniques are also applicable to synthesizing safety data in the reference population(s) and consider adding the following sentence: Data on other drugs with same or similar MoAs can be informative as well (when pharmacologically contextualized and data permitting), and integrated using model-based meta-analyses	Consider adding the following sentence: Data on other drugs with same or similar MoAs can be informative as well (when pharmacologically contextualized and data permitting), and integrated using model-based meta-analyses.
EFPIA	452	453	3,6	"Meta-analytic techniques for synthesizing efficacy data 452 in the reference population(s) should also be considered." Use of meta-analytical techniques is proposed for efficacy data only. Because the guideline emphasises extrapolation of efficacy and safety the latter should be also subject of a meta-analysis.	Rephrase sentence "Meta-analytic techniques for synthesizing efficacy and/or safety data in the reference population(s) should also be considered."

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	455	465	3,6	Bayesian approaches to evaluate similarity of response between the reference and target populations should be included in the text, highlighting that they can be very appropriate because, as the guideline states (lines 237-238), "similarities and differences of response to treatment between a reference and target populations should be understood as a continuum". Moreover the frequentist approach would fail in the event of increased uncertainty: the use of Bayesian approaches can accommodate this and leverage as much information from adults as possible.	Frequentist approaches to evaluate similarity of response between the reference and target populations can be informed by a comparison of point estimates and their associated confidence intervals. Given the different levels of precision typically available for estimating parameters in different populations, it will often be inappropriate to declare similarity purely based on overlapping confidence intervals. Bayesian hierarchical models and dynamic borrowing techniques such as commensurate priors for example, could be used to integrate and synthesize available evidence into a probability distribution, when the degree of similarity (exchangeability) is in a continuum, avoiding the need of a binary yes/no decision to the similarity assessment.
EFPIA	456	457	3,6	"... the parameters being evaluated for similarity" this comes out of the blue. What do you mean here?	
Lundbeck	457	459	3.6	It seems to be suggested above that similarity of response is also a continuum. Is it correctly understood that sometimes extrapolation could for example refer to that if there is an effect in adults it can be assumed that there is an effect in children, without the effect size being exactly the same? The exact effect in children may therefore need to be investigated. Can the effect size be understood based on modeling or would that require a trial and if it requires a trial, how then to design it as it may not need to be powered for detecting an effect, but rather to obtain a certain precision? Could you please elaborate?	
EFPIA	457	459	3,6	Beside frequentist approaches also Bayesian approaches can be used to assess similarities by comaring posterior distributions and/or characteristics of posterior distributions like mean, median, sd, credible intervals and quantiles which also gives more insided on difference than the frequentist approaches. Further, distributions of patient endpoints and/or aggregated patient endpoint data give further insight on the similarities.	Add Bayesian approaches to assess similarities based on posterior distributions or its characteristics. In addition, predictions of endpoints or aggregated endpoint data, derived e.g. by using posterior predictive distributions should be added as valuable tool to assess similarities or differences.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	457	457	3,6	The word "parameter" may be misleading here. Probably, "endpoint", or "outcome" would be more appropriate. It would be good if the guidance would include a glossary of terms, all technical terms should be defined, and then used in a harmonized manner throughout this document. Someone needs to go through this document and polish it. It is very noticeable that different authors have written the text, without agreement on terms.	
EFPIA	457	462	3,6	This entire part seems out of context. To focus on 1 very specific approach to 1 small component of methodology in the setting of a more high-level section seems strange to me. Either touch briefly on a variety of different approaches here (without getting technical about precision and overlapping confidence intervals) or remove this part altogether and discuss different approaches in a dedicated section.	
EFPIA	459	464	3,6	This text jumps from confidence intervals to a model. What model are you talking about here? What is appropriate if overlapping confidence intervals are inappropriate? What does the last sentence starting with "Communication of the ..." really mean? It doesn't seem to have too much meaning in this context.	
EFPIA	461	462	3,6	The stated premise that overlapping CIs are not sufficient lacks a viable counterpart. Given that similarity is a continuum, the term "declare similarity" is not a well defined term here. Clarification of the intention is requested. Furthermore, for clarity of wording it is suggested replacing the word "declare".	A formal equivalence test with an informed equivalence bound on the other hand could constitute a viable strategy for assessing similarity. Given the different levels of precision typically available for estimating parameters in different populations, it will often be inappropriate to establish declare similarity purely based on overlapping confidence intervals.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Agios	462	465	3,6	Extrapolation can often be more meaningfully performed using Bayesian approaches than frequentist approaches.	Communication of the manner in which uncertainty has been defined, specified, and otherwise accounted for in the model development and any simulations used to assess similarity of disease and/or response is recommended. In addition, any relevant assumptions with respect to the definition or expression of uncertainty should be specified. Bayesian approaches can also be used and require full specification for the model including prior distributions for model parameters.
Takeda	462	464	3,6	Communication with whom is unclear.	
Koop Phyto	464	464	3.5.2	see line 19	similarities of disease, response to therapy, adverse reactions and undesired events
ISCT	473	476		Sponsors must discuss acceptability of the proposed approach with regulatory authorities	
Koop Phyto	473	473	3.5.2	see line 19	similarities of disease, response to therapy, adverse reactions and undesired events

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	473	475	3,6	What is meant by "uncertainty in the data" here? The fact that in the vast majority of the cases, at time of defining extrapolation concept, there is no response data for the investigational drug in the target population/indication, and thus assessment of similarity involves untestable assumptions?	
ISCT	478	484		Gaps in knowledge should be addressed prior to paediatric extrapolation concept being finalised.	
Lundbeck	479	481	3.6	Could you clarify what kind of gaps may need to be addressed prior to finalizing the concept? It was our understanding that the extrapolation plan was intended to address the gaps? (see also Comment 3)	
EFPIA	488	489	3,6	Please include recommendations about when the extrapolation concept can be finalized, in spite of remaining gaps, or at least provide examples.	
EFPIA	488	490	3,6	We wonder in which case gaps would need to be addressed for the extrapolation concept itself. Isn't the purpose of the concept to identify gaps and the plan to collect data to fill in the gaps?	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	488	490		It would be helpful to learn about examples of knowledge gaps which will be not part of the plan but needs to be completed before the extrapolation concept can be finalized. Some regulators expect ped development plans early in the adult development – how can this fit together?	
EFPIA	496	497	3,6	It is suggested that the word "Presentation" be clarified.	3.7 Outline Presentation of the Pediatric Extrapolation Concept
EFPIA	513	513	4	Could an example template be provided for the Extrapolation Framework?	
EFPIA	513	519	4	This section on the pediatric extrapolation plan implies that the only way to support the extrapolation concept is to generate new data via relevant study(ies). This is most likely what will need to happen, however it is not always the case. They are a few examples where extrapolations to a specific pediatric population were approved only based on simulations and without collecting data such as the following example: https://www.page-meeting.org/?abstract=9089	It is suggested adding a paragraph indicating that when the data are difficult to collect, extrapolations could be potentially based only on simulations. This would mean of course that the model is robust, that the uncertainties associated with the data supporting extrapolation to the target pediatric population are limited and that the expected benefit outweigh the risk.
EFPIA	518	519	4	The current text may seem too restrictive. Suggest rewording. Moreover, It would be beneficial for sponsors to consider discussing the acceptability of the extrapolation plan with regulatory authorities. Suggest rewording.	"The study elements in the pediatric extrapolation plan, such as the design, timing, analysis, interpretation and reporting of studies, included in the pediatric extrapolation plan are elaborated below. Sponsors should consider discussing the acceptability of the extrapolation plan with regulatory"

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	522	557	4,1	This entire section contains a range of very good and valuable points, but it lacks a clear structure and hence needs considerable improvement. As an option (1) Start with the objective: An efficacious dose in the target population is often determined as a dose that achieves the same exposure that has been demonstrated to be safe and effective in the reference population (exposure matching). (2) Explain the assumptions that justify this objective: a key assumption here is similar exposure-response curve between reference and target population. (3) Explain the type of evidence that is required to support the assumption: establishment of E-R relationships in reference and target population would provide sufficient evidence (if similar), but this is not always possible and is hence not required by this guidance. For example, because randomization to sub-therapeutic or supra-therapeutic doses in the pediatric target population may be unethical. Often one will only test one pediatric dose in the pediatric target population. If one can demonstrate that this dose achieves similar exposure in the target and the reference population, and similar efficacy, then this is sufficient evidence. (4) Explain what is needed to establish that the initial dose estimate meets its objectives (usually the objective is to find a dose that matches the adult exposure). This is usually done by conducting a dedicated pediatric PK study or a PK cohort as part of a dedicated pediatric efficacy/safety study and by demonstrating that the observed exposure indeed meets the objective.	
EFPIA	527	528	4,1	"As part of planning for dose selection, other considerations (e.g., safety, 527 formulation, final dosing regimen) should be incorporated." Formulation is an important part in pediatric drug development as often pediatric formulations need to be developed.	More detailed guidance addressing formulation considerations would be welcome (for example whether a full bioequivalence assessment is required and in which population to allow extrapolation) but also to strengthen the need to develop the right formulation and the right device for children.
ISCT	534	620	4,1	No comment - outside area of expertise	
Agios	534	536	4,1	Should read "relationship" rather than "curve"	It is important to note that the identification of safe and effective dose(s) in the program with the reference population does not always require or result in the demonstration of an exposure response (E-R) relationship. As such, there is no requirement to establish an E-R relationship in pediatrics.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
ISCT	536	536	4,1	Cannot make a conclusion or comment on the validity of the statement 'As such, there is no requirement to establish an E-R curve in pediatrics.', but suggest that the statement is examined by those with appropriate expertise to determine if it needs adjustment or caveating.	
ProPharma Group	537	539	4,1	Previous Agency feedback received confirmed that only when the dose-exposure-efficacy relationship in adults is clear, can this be used for extrapolation of efficacy from adults to to paediatric population. Suggest it is made clear that the timing is critical.	n/a
EFPIA	537	542	4,1	The section is hard to understand. It is over complicated with multiple negations (lack of X does not preclude Y). Why not rather simply state: "Exposure-matching may still be utilised in the absence of demonstrable E-R relationships when the expectation that a comparable response at the target drug exposure is likely to be achieved"	Why not rather simply state: "Exposure-matching may still be utilised in the absence of demonstrable E-R relationships when the expectation that a comparable response at the target drug exposure is likely to be achieved"
EFPIA	537	539	4,1	"However, the lack of demonstrable E-R relationship in the reference population or the inability to demonstrate similar E-R curves in the reference and target populations does not preclude the use of exposure matching for dose selection purposes in the pediatric extrapolation plan". Suggest amending "E-R curves" to "E-R relationship", for clarity. This is not a "However"; lack of E-R does not preclude exposure matching. Correct syntax. Please delete the word 'However," and simply start the sentence with "The lack of...". This is an adverb that's supposed to introduce a statement that contrast with the previous one. This is not the case here and could be misleading.	THowever, the lack of demonstrable E-R relationship in the reference population or the inability to demonstrate similar E-R relationship curves in the reference and target population …
EFPIA	539	542	4,1	"Dose selection based on exposure matching under such circumstances is reasonable and pragmatic and is predicated on the expectation that comparable response at the target drug response is likely to be achieved." This sentence should be modified as it does not make sense in the context of exposure matching.	"Dose selection based on exposure matching under such circumstances is reasonable and pragmatic and is predicated on the expectation that comparable response at the target drug response exposure is likely to be achieved."
Agios	542	543	4,1	Should also call out cases where it may be unethical to randomize pediatric patients to a non-active control (placebo)	Furthermore, there are situations in which randomization of pediatric patients to a non-active control (ie, placebo) may be unethical

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	542	544	4,1	It is specified that there are situations where it may be unethical to randomise paediatric patients to subtherapeutic doses while the safety data available do not support evaluation of higher doses/exposures. Please provide further guidance or more specifics for such example(s). What happens in such situations that is not currentlyl clear from the guidance and does the paediatric development stop then or how is this issue to be overcome?	
GPT	545	557		We appreciate the differentiated way that is used here to distinguish between PK data necessities and situations that do well without, because this reflects the different degree of relevance.	
EFPIA	545	557	4,1	Please define what is meant by "confirmatory PK". This section contains many good points, but it is not clearly structured. Please see below suggestions for improvement. We guess what this section is trying to say is that the initial pediatric dose will usually be determined based on data from the reference population. The objective is to find a dose regimen in the target pediatric population with similar exposure as in the reference population (exposure matching). Once determined, evidence should be provided to show that the objective was met (i.e., that this dose indeed matches the reference exposure). This can be done with a pediatric PK study, but in many circumstances, it suffices to demonstrate exposure-matching as part of a pediatric efficacy/safety study. If pediatric PK data from same drug but different disease are available, these could also be used.	
EFPIA	548	553	4,1	The obligation to conduct a PK study may at times be unethical e.g. explore existence of non linearity in children and age categories. As such, it is recommended that it be made clear that the requirement to conduct this be determined on a case by case basis. An alternative approach using modeling and simulation might also be considered. Other scenarios where a separate PK study be executed should be considered i.e. when PK extrapolation from reference to target population is expected to fail. It is indicated that a separate PK study should be conducted in certain situation, however an alternative could be to have a sub-PK study with a richer PK sampling design embedded in an efficacy/safety study. See proposed revisions.	However, a separate PK study should be considered, where appropriate, conducted in certain situations (e.g., drugs with narrow therapeutic range, linear PK, and/or potential differences in the effect of disease on the PK of the drug between the reference and target populations, and/or when PK extrapolation from reference to target population is expected to fail). Alternatively, consideration may be given to using modeling and simulation approaches such as physiologically based pharmacokinetic (PBPK) approaches. Moreover, if richer PK information than sparse PK is needed then it could also be collected in a sub-PK study part of a larger clinical study.
EFPIA	550	553	4,1	Please explain why one can't use PK information obtained from an efficacy/safety conducted in the target pediatric population here? In case of "differences in the effect of PK of the drug between reference and target population", is exposure-matching still a good objective?	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	553	557	4,1	"Lastly, when PK data are available in an adult reference population with the disease and the exposure is within an observed exposure range in a reference pediatric population with a different disease(s), additional PK assessment may not be necessary in the target population" It is a bit convoluted, see proposed rewording. OR The reference pediatric population with a difference disease and target population should generally be with same age range, (e.g., both populations are 1~5 years old). Otherwise, it is not applicable to believe that PK assessments are unnecessary. Please revise the sentence to reflect this. It is proposed to re-write the guidance text .	Lastly, when observed exposure PK data are available in an adult reference pediatric population with the disease and the exposure is within an observed exposure the range of PK data available in an adult reference pediatric population with a different disease(s), additional PK assessment may not be necessary in the target population. Lastly, additional PK data in the target pediatric population may not be required if there are PK data on the experimental drug from a different pediatric population / indication. These data should usually include the same age range as relevant for the target pediatric indication. when PK data are available in an adult reference population with the disease and the exposure is within an observed exposure range in a reference pediatric population with a different disease(s), additional PK assessment may not be necessary in the target population; Hhowever, this approach relies on understanding the effect of disease on the PK of the
EFPIA	559	559	4.1.1	The header is not a question but has a questionsmark. Please rephrase.	When Should Dose Ranging Data Should be Collected?
EFPIA	559	562	4.1.1.	Dose-ranging studies in pediatrics are often challenging. Due to the small sample sizes of such trials, it often renders inconclusive or uninterpretable results. It is suggested that pediatric dose ranging studies should only be conducted if there is large uncertainty in disease similarity/prior evidence of disease dissimilarity between adults and pediatrics.	Dose ranging data may be needed as part of the pediatric extrapolation plan. Such circumstances may include when there is large uncertainty in the disease similarity and/or response to treatment (exposure-response); when there are potential …
EFPIA	564	564	4.1.1.	Clarification is requested as to whether this is intended to be D-R or E-R.	
EFPIA	569	569	4.1.2	Use of biomarkers seems to give general guidance regarding biomarker endpoints rather than dose selection specific guidance (it is part of 4.1 dose selection).	It is suggested to move 4.1.2 to a standalone biomarker section or clarify.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	574	576	4.1.2	Can the guidance provide some examples on how PBPK or QSP may be used to support biomarker or clinical endpoint in children? PBPK is not relevant for biomarkers, it is really PK focused. It is suggested to replace by PK/PD. Not sure how PBPK models can help with mechanistic representation of biomarker time course data in response to drug therapy. Although in principle PBPK models can incorporate population specific mechanisms that allow prediction of target organ concentrations, in the absence of validation, these resource intensive models can be misleading and lead to unnecessary delays.	
EFPIA	578	579	4.1.2	A biomarker may or may not need to be validated, although use of a validated biomarker may require less justification.' How much is this aligned with other situations in adult development that a biomarker potentially used as primary endpoint may not need to be validated? The guideline states that a biomarker may or may not need to be validated in the context of dose selection and the use of a validated biomarker as a surrogate endpoint is recommended but not required. However, the guideline is silent on what level of evidence is needed to justify the use of a biomarker that is not validated.	Recommend including guidance on what level of evidence regulatory authorities require when sponsors choose to use a biomarker that is not validated in support of dose selection or when establishing efficacy. Alternatively, provide examples how a non-validated biomarker can provide evidence to support dose selection or establish efficacy and include strategies around interpretation of data.
EFPIA	579	579	4.1.2	Methodological considerations (e.g., the effect of missing data, and the results of sensitivity analyses to departures from any assumptions) should also be included in the evaluation of the proposed endpoint [see ICH E9(R1)]'. Mention definition of estimands	… (e.g. definition of primary or secondary estimands, the effects of missing values, …)…
Lundbeck	581	581	4.1.2	1) Is there a specific reason ICH E9(R1) is mentioned only in this biomarker section? 2) The guideline is overall quiet about any recommendations on the use of estimands in pediatric studies, are there any considerations regarding estimands and extrapolation?	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	583	586	4.1.2	If a biomarker has been proposed for use as a primary analysis in the target population and cannot be measured in the reference population, relevant clinical outcomes in the target population should at least be measured as well, to try and understand the relationship between the variables.' Not clear what would be the value of such a biomarker as primary endpoint in the context of extrapolation from the reference (i.e., in the context of this guidance). It is impossible to extrapolate from reference to target population on the basis of a biomarker that has not been measured in the reference population. If it is not for extrapolation, maybe you should mention that such a study is not covered by the guidance. It may be that the primary endpoint can be measured in the reference population, but simply was not measured for any reason. Moreover, it is suggested to be more precise ('between the variable).	Proposed change: If a biomarker has been proposed for use as a primary analysis in the target population and cannot be or was not measured in the reference population,.... between the variables biomarker and the clinical outcome.
EFPIA	588	589	4.1.3	Scenarios for dose selection' The concept of generalization is missing in the list - where no additional pediatric data is needed to be collected.	
EFPIA	590	723	4.1.3.1, 4.1.3.2, 4.1.3.2	It seems inappropriate to list these subsections under Section 4.1.3 "Scenarios for Dose Selection", since each of these subsections corresponds to appropriate considerations for PK study designs under different scenarios. Although dose selection is important in these design options, these subsections contain much more information than dose selection. In addition, dose selection is also important for efficacy studies, therefore the flow and section organization can be improved for this part.	Suggest modifying the title of the section: 4.21.3 Scenarios for Dose Selection PK Studies, and updating each subsection and other sections accordingly.
EFPIA	590	590	4.1.3.1	When only PK data are Needed to Establish Efficacy What if the outcome of the study is negative (exposure in the target population does not match that in the reference population)? Should we redo a new PK study or is it acceptable to determine the paediatric regimen by modelling and simulation based on this “failed” paediatric study? If the latter is possible, it should be mentioned in the guidance. Please clarify.	
EFPIA	594	605	4.1.3.1	Please clearly define the dosing strategy based on exposure matching. Does it always consist of selecting a paediatric regimen that achieves an exposure in the target population similar in mean and distribution to that in the reference population treated with the approved regimen? Any exception to this rule, e.g., when the adult bodyweight extents much beyond the paediatric bodyweights, and thus the matching exposure strategy may result in paediatric patients receiving a lower dose compared to adults of same bodyweight?	
EFPIA	596	596	4.1.3.1	"Modeling and simulation strategies should be applied to support the initial dose selection in the exposure matching study in the target population." Please define 'exposure-matching study' and clarify if it means a Phase 1 single-dose study, or can other Phase 2/3 studies provide this information.	The guideline should define 'exposure-matching study' and clarify whether it means a phase 1 single-dose study or if other phase 2/3 studies can provide the necessary information.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	603	605	4.1.3.1	Wouldn't it also suffice to demonstrate empirically that the exposures in the two populations match? What the guidance should really discuss here is the following: (1)Based on adult reference data and a PK model one estimates an initial pediatric dose. Usually, the dose that is being proposed is a dose that when applied to children matches the exposure seen in adults. (2)This initial dose is tested, either in a stand alone PK study in a PK cohort that is part of a pediatric efficacy/safety study. (3)The resulting PK data from that study or cohort should be assessed to see whether it indeed matches the exposure seen in adults. (This assessment can be by direct comparison, or by model-based comparison. Direct is preferrable, as it does not rely on assumptions. The guidance should explain under which circumstances a model-based comparison should be used). (4)If the exposure obtained with the initial pediatric dose doesn't match the adult exposure, the PK model should be updated, a new pediatric dose should be proposed and then usually tested in a PK study or cohort again. This should happen as soon as possible. One should not wait until the end of a pediatric efficacy/safety study to assess PK, but rather perform an interim analysis. Therefore, any pediatric development plan should include a certain "adaptive" element to cover this aspect of having to re-estimate the pediatric dose. (5)The study design section of this guidance should comment on this "adaptive" approach as well.	
EFPIA	603	605	4.1.3.1	It needs to be pointed out that the final dosing regimen for labeling does not only depend on PK simulation.	Suggest changing the sentence as follows: "the proposed dosing regimen should be re-evaluated through simulation techniques along with other evaluations (e.g. safety extrapolation if needed) before a final dosing regimen for proposed product labeling is selected."
EFPIA	604	604	4.1.3.1	Once PK data are obtained in the target population, the proposed dosing regimen should be re-evaluated through simulation techniques before a final dosing regimen for proposed product labeling is selected.' see proposed revision.	… through modeling and simulation techniques…
EFPIA	606	626	4.1.3.1	Model-informed dose selection feasibility and practicality of dosing strategies as well as the sample size feasibility is mentioned but the guideline does not include further details on the feasibility. There is currently limited information or guidance in this document for when there are situations or programs where there is lack of prior data available, lack of validated pediatric endpoints or sample sizes, and circumstances when efficacy studies are not required.	It would be beneficial if the feasibility topic briefly mentioned in the guideline was expanded. Additional guidance and recommendations regarding feasibility in general for pediatric studies as well as the special situations mentioned (lack of prior data, lack of pediatric validated endpoints, etc.) would be helpful when developing pediatric programs.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	607	607	4.1.3.1	Endpoint: Target exposure metric' Is it necessary to design a PK study in the target population allowing to evaluate the steady state PK metric or could it be acceptable to characterize the PK based on a single dose study (and extrapolate the steady state PK metric based on modelling)? If this is possible, it should be mentioned in the guidance.	
EFPIA	608	608	4.1.3.1	When the pediatric extrapolation strategy relies on matching adult exposures, the target exposure metric(s), range, and acceptance criteria should be prospectively specified and should be defined in the context of the disease, treatment regimen, route of administration, and formulation.' The wording "matching adult exposures" could be clarified. Why only "adult exposures" and not exposures in the reference population? It is worth clarifying that "matching adult exposures" means that the proposed pediatric dose produces exposures that are within the effective and safe exposure ranges based on adult data, and does not necessarily mean that the pediatric exposure profile exactly match the observbed adult exposure profile. Does this only apply to matching to adult exposures, or should this statement be about matching exposures from a reference population (which could be adults, or perhaps adolescents or another relevant pediatric population?). The rest of the paragraph is more generic and refers to reference population not just adults. It would be difficult to pre-emptively specify a criteria for an 'acceptable' exposure range, as subsequent pediatric studies might demonstrate a slightly different exposure-response relationship than that observed in the reference population, which would result in evaluation/approval of a slightly different exposure range and therefore, dose regimen. See proposed revisions.	When the pediatric extrapolation strategy relies on matching adult exposures from the reference population, the target exposure metric(s), range, and acceptance criteria should may be prospectively specified if applicable, and should be defined in the context of the disease, treatment regimen, route of administration, and formulation.
German Pharmaceutical Industry Association	622	642	4	The clear guidance on sample size justification is appreciated.	
AESGP	622	642	4	The clear guidance on sample size justification is appreciated.	
EFPIA	622	622	4.1.3.1.	The ‘criterion for success’ is imprecisely defined in the two sections. It is understandable that defining such a criterion is not simple as it depends on many aspects, including feasibility, and width of therapeutic range. Still, it is important to define it clearly and concisely in the sample size section.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	622	698	4.1.3.1	This section is too imprecise and unclear. It needs a complete re-writing, which should be guided by practical considerations.	
EFPIA	622	622	4.1.3.1	As sample size is an end product from evaluating how likely the trial will meet the criteria (such as passing a pre-specified threshold), it would provide more guideline to emphasize the importance of the pre-specified criteria in the extrapolation plan	Suggest changing the section title from "Sample Size" to "Criteria and Sample Size". This suggestion applies to Line 693 as well. Also it would be helpful to include more guidance or examples on when precision study design and when matching the pre-specified target exposure range should be considered
EFPIA	622	642	4.1.3.1	Optimal design and clinical trial simulation can provide guidance on PK sample, however practical considerations also play a role. PK sample schedule suggested by optimal design may not be practical in all cases. Can the language in lines 639-640 be softened so that this would not be a requirement?	
EFPIA	622	622	4.1.3.1	Can we advise on a kind of minimum sample size?	Except in rare disease where recruitment would be very difficult, a sample size below 10 patients would not be adequate.
EFPIA	624	625	4.1.3.1	One should expect a more substantiated recommendation. Simply stating that Sponsors should have adequate representation is not enough information. The guidance should provide further elements of rationale. For example, if the authors have an opinion if we should use age subgroups or weight subgroup, then they must have an idea about how to create the subgroups. And then we would expect the recommendation to recruit paediatric patients according to the (age or weight) distribution in the paediatric population but acknowledge some limitations and so on.	
EFPIA	624	624	4.1.3.1	"The sample size for a pediatric PK study should be sufficient to meet the objectives of the study and be based on quantitative methods (modeling and simulation and/or statistical approaches)". "modelling and simulation" are also statistical approaches, therefore it may not be appropriate to separate them from "statistical approaches". Also, this sentence should be revised since quantitative methods may not always apply. Most PK studies are conducted with a sample size selected based on feasibility, for purposes if empirical comparisons also supporting population PK/PD analyses, and ethical considerations. What if the sample size can only be based on feasibility aspects? See proposed revisions.	"The sample size for a pediatric PK study should be sufficient to meet the objectives of the study and may be based on quantitative methods (modeling and simulation and/or other statistical approaches), as appropriate"

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	629	637	4.1.3.1	The sample size justification and its feasibility in the targeted indication should include the following: . The methodology(ies) used to determine the sample size, including the decision criteria [This should come first]. In addition, minor clarification proposed on lines 631-632: The adequacy of the sample size to estimate the key PK parameters in the pediatric population (such as clearance and volume of distribution) with adequate precision" and recommend to provide further guidance in what constitutes an "adequate precision". The availability of pediatric data from other disease population can be useful; consider adding a bullet point as proposed.	•The methodology(ies) used to determine the sample size, including the decision criteria •The availability of patients in a specific body weight/age range • The availability of pediatric data from other disease populations •The adequacy of the sample size to estimate the demonstrate precision in key PK parameters in the pediatric population (such as clearance and volume of distribution) with adequate precision •The adequacy of the sample size to match the pre-specified target exposure range (e.g., the interquartile range for the PK metric(s) in the reference population) •The methodology(ies) used to determine the sample size
EFPIA	631	632	4.1.3.1	The adequacy of the sample size to demonstrate precision in key PK parameters in the pediatric population such as clearance and volume of distribution'. Parameters like clearance and volume are determined based on a popPK model. This popPK model will usually be based on a pooled data set which may include adult patients and patients from other diseases treated with the same drug. The sample size for such a popPK model should be seen as a separate problem to the sample size for a pediatric PK study to demonstrate exposure matching. This guidance should make this clearer to the reader. Most of the time the pediatric PK data are pooled with adult PK data and analysed using a population approach. In such type of analysis, the pediatric sample size as a limited impact on the precisions of the population CL and V values due to the informative adult dataset. However, the parameters that could be sensitive to the pediatric sample size are the parameters characterizing the expected covariate effects on CL and V such as body size and age. Those parameters would also be sensitive to the expected distribution of those covariate in the pediatric sample size.	It is suggested adding that the sample size calculation should include the adequacy of the sample size AND the distribution of the expected pediatric covariate to demonstrate precision in the parameters characterizing those expected pediatric covariate effects.
Pharmetheus AB	634	635		This sounds like an equivalence criteria and is not in line with the requirement to estimate the precision in the PK parameters.	See Comment 1 to section 4.1.3.1

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	634	635	4.1.3.1	The adequacy of the sample size to match the pre-specified target exposure range (e.g., the interquartile range for the PK metric(s) in the reference population)'. The adequacy to match a range is not very clear and accurate, so rephrasing is recommended. It is also suggested that line 634 be amended as shown (for example, for PK it could mean that sample size is based on being able to detect CL changes that would lead to a different dose selection).	• The adequacy of the sample size to match the pre-specified target exposure range (e.g.,the interquartile range for the PK metric(s) in the reference population) detect if key assumptions made for dose selection or other study characteristics deviate enough to change prior conclusions.
EFPIA	639	640	4.1.3.1	Modeling and simulation techniques such as optimal design and/or clinical trial simulation should be conducted to justify the timing and number of PK samples.' This seems too strong, suggest revision.	Modeling and simulation techniques such as optimal design and/or clinical trial simulation should can be conducted to justify the timing and number of PK samples.
German Pharmaceutical Industry Association	644	669	4	The guidance on analysis and reporting is highly appreciated.	
AESGP	644	669	4	The guidance on analysis and reporting is highly appreciated.	
EFPIA	645	669	4.1.3.1	This subsection (Analysis and Reporting) almost contains two different and contradicting philosophies. There are some statements (line 648-650 or lines 664-669) which recommend using simple methods like confidence intervals or graphical procedures. There are other parts (654-659) where model-based approaches are recommended. We agree very much with the simple approach, which should be used if possible. If not, one should use a more complex approach. What is missing is a discussion under which circumstances to use the simpler approach, and under which circumstances to use a model-based approach. What is also missing is a high-level description of how such a model-based approach could look like.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	645	652	4.1.3.1	"Analysis and reporting Different presentations of the exposure data in the target and reference populations should be available to inform regulatory decision making. 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 confidence intervals for the mean differences in key exposure metrics such as AUC and Cmax could be an acceptable approach. The chosen boundaries 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 single acceptance boundary for all drug 646 products and drug classes (as compared to bioequivalence testing) will not provide a 647 meaningful approach in the setting of pediatric extrapolation': this statement seems to assume that one could do some sort of equivalence test on exposure data. First of all, the approach should be described before commenting on it. Second, are you seriously proposing an equivalence test? Between what? Adult exposure data and pediatric exposure data (which would come from different studies)? That would potentially require a huge sample size, unless acceptance boundaries are so wide that the approach is meaningless anyway. Or is the proposal to compare predictions based on a model with observations in the target population? Please clarify. This section should also note that there may be many instances where exact exposure-matching is not warranted for selection of the most appropriate pediatric dosage, such as when target expression in pediatrics can differ, or when the dose can be selected primarily from PD response matching instead. See proposed edits	Analysis and reporting Different presentations of the exposure data in the target and reference populations should be available to inform regulatory decision making. 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 confidence intervals for the geometric mean differences in key exposure metrics such as AUC and Cmax could be an acceptable approach. The chosen boundaries 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. However, there may be many instances where exact exposure-matching is not warranted for selection of the most appropriate pediatric dosage, such as when target expression in pediatrics can differ, or when the dose can be selected primarily from PD response matching instead.
ISCT	646	964	4.1 - 4.3	No comment - outside area of expertise	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Pharmetheus AB	648	652	4.1.3.1	"An evaluation of confidence intervals for the mean differences in key exposure metrics such as AUC and Cmax could be an acceptable approach. The chosen boundaries 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." It needs to be clarified, what the primary PK endpoint in a PK study is; more information is needed here.	Comment 1: We interpret this sentence as if an approach similar to bioequivenence is required, which we believe is a step in the wrong direction. Previous guidelines have focused on that the PK parameters should be estimated with sufficient precision, which we hope also will be the case in this guideline. For example, in the case of extrapolation from adults to the pediatric population using PK: a dose regimen to be used in a pedatric study is proposed based on the best available information (PK models and target value for key exposure metric), and, the sample size is estimated with the purpose to obtain sufficient precision in the primary PK parameter, usually CL. The study is performed, and PK is evaluated. Provided that the primary PK parameters (CL) is estimated with sufficient precision, the dose regimen may be adjusted, if needed, based on PK predictions. Thus, in this example, the primary PK parameter would be CL, ie not the key exposure metric. Comment 2: As stated in this guideline, the appropriate precision boundaries should be guided by the therapeutic window. In addition, it could be mentioned that the lack of clinical impact of these boundaries may be demonstrated through simulations if an exposure-response curve has been established in the reference population.
EFPIA	650	652	4.1.3.1	The chosen boundaries 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.' Are you talking about the threshold (e.g., 0.8, 1.25) or the significance level of the confidence interval? The later could be informed by the risk benefit but not by the therapeutic range. vice-versa for the former. Please clarify.	
EFPIA	654	657	4,1	A model-based comparison (that can integrate all available data) is generally preferred rather than a descriptive comparison of observed adult and pediatric exposure data alone. In addition, inter-individual variability needs to be considered in establishing exposure similarity rather than comparing means alone.' What kind of model-based comparison? Please explain. Moreover, a direct comparison of observed exposure data in reference and target population by graphical means is preferrable, because it does not rely on any assumptions. So I would suggest change preference and to explain WHEN model-based comparisons might be preferrable over direct ones.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	657	659	4.1.3.1	"A simulation of the percent of subjects at different age/weight ranges that lie within (or outside) a pre-defined exposure range may provide a more meaningful assessment of exposure similarity."	A simulation of the percent of subjects at different age/weight ranges that lie within (or outside) a pre-defined target exposure range may provide a more
EFPIA	661	663	4.1.3.1	In general, the most relevant covariate to influence PK in pediatric patients is body weight. In the youngest pediatric patients (e.g., infants and neonates), in addition to body weight, age is also an important covariate to account for relevant organ maturation.' We would suggest to consider specifying age ranges where maturation typically applies and specifying that functions should be ideally be continous. For more clarity, we suggest providing the definition of neonates (ie, 0-28 days old) and infants (>28 days to 1 year).	In general, the most relevant covariate to influence PK in pediatric patients is body weight. In the youngest pediatric patients ([e.g., infants (>28 days to 1 year) and neonates (0-28 days)]), in addition to body weight, age is also an important covariate to account for relevant organ maturation.
EFPIA	664	665	4.1.3.1	Relevant predefined exposure metrics should be presented graphically versus body weight and/or age on a continuous scale.' Delete "on a continuous scale". In some cases, it would be better to treat age and BW as use category covariates.	Relevant predefined exposure metrics should be presented graphically versus body weight and/or age on a continuous scale.
Lundbeck	670	711	4.1.3.2	Biomarkers are discussed in several places in this section on dose finding, but the considerations seem to be quite general and also relevant for the efficacy trials discussed in section 4.3. It would be helpful if the document could be restructured regarding how biomarkers are discussed. The current setup almost suggests that biomarkers can only be used with regards to dose-finding, but we get the impression here that they may also be used to establish efficacy?	
EFPIA	670	691	4.1.3.2	When Effect on a Biomarker is Needed to Establish Efficacy.' More thought is needed for this important subsection. In principle there are two ways how biomarkers can be used. One could use an exposure-biomarker relationship for extrapolation (i.e. the biomarker replaces the response), or one could use a biomarker matching strategy (i.e. the biomarker replaces exposure, and one needs to find a pediatric dose that matches the biomarker levels observed in the reference population). These different approaches should be discussed in more detail.	
Takeda	674	674	4.1.3.2	Are there any definition for "validated biomarker"?	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	674	674	4.1.3.2	Use of a validated biomarker as a surrogate endpoint is recommended but not required.' What is meant by "Use of a validated biomarker as a surrogate endpoint is recommended but not required". The biomarker does not need to be validated or other endpoints can be used? The guideline states that a biomarker may or may not need to be validated in the context of dose selection and the use of a validated biomarker as a surrogate endpoint is recommended but not required. However, the guideline is silent on what level of evidence is needed to justify the use of a biomarker that is not validated. Please clarify.	Recommend including guidance on what level of evidence regulatory authorities require when sponsors choose to use a biomarker that is not validated in support of dose selection or when establishing efficacy. Alternatively, provide examples how a non-validated biomarker can provide evidence to support dose selection or establish efficacy and include strategies around interpretation of data.
EFPIA	676	677		The choice of the biomarker endpoint should be supported by available data in the reference and target populations and justified in the extrapolation plan.' This does not seems fully consistent with the previous section 583-586. please reword.	The choice of the biomarker endpoint should be supported by available data in the reference and or target populations and justified in the extrapolation plan.
EFPIA	679	681	4.1.3.2	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.' This bullet would better fit as second bullet, consider moving it up (Line 676). Indeed it may a good way to mitigate the absence of a surrogate endpoint which is usually quite demanding. In addition, we would rather suggest associated rather than correlated to not confuse with statistical concept underlined with "correlated".	A biomarker on the causal pathway that is correlated associated with clinical efficacy in the reference population is often acceptable and should be justified also with regard to its relevance to the target population.
EFPIA	686	691	4.1.3.2	Biomarker in the target (ped) population may also be used as the only feasible proxy as endpoints are too rare and population sizes too small. In that case a biomarker maybe used even without proven relationship to efficacy in the reference population. Further some ped biomarker may not be measured/measurable in the adult population and eventually anyhow the relationship in the TARGET population is of relevance (see also 583-586). Please define what could establish "confidence" in a relationship between biomarker and efficacy.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
German Pharmaceutical Industry Association	693	711	4	The guidance on sample size, analysis and reporting for biomarker trials is also very helpful.	
Takeda	693	698	4.1.3.2	Should feasibility not be considered to set the sample size, too?	
AESGP	693	711	4	The guidance on sample size, analysis and reporting for biomarker trials is also very helpful.	
EFPIA	694	695		Quantitative methods (modeling and simulation or statistical approaches) should be used to derive sample size for PK/biomarker and biomarker endpoints.' There is probably a difference whether the sample size is based on PK or on biomarker outcome. Also, consider enlarging here to dose.	Quantitative methods (modeling and simulation or statistical approaches) should be used to derive sample size for PK/biomarker and biomarker endpoints.
EFPIA	695	696	4.1.3.2	The Sample Size section is not being handled consistently for the different scenari making unclear what is expected depending on the respective situations. In general, the sample size must be justified and guided by the decision criteria	Replace paragraph on sample size with : The sample size justification and its feasibility in the targeted indication should include the following: * The methodology(ies) used to determine the sample size, including the decision criteria [This should come first] * The availability of patients * Consideration of the bio-sampling feasibility (number of biomarkers, timing/frequency, volume, ...) * The variability of the key biomarkers relevant for the decision
EFPIA	703	705	4.1.3.2	A therapeutic range of the biomarker effect that provides a meaningful assessment of similarity between the reference and target 'populations should be pre-defined.' The wording "pre-defined" needs clarification. It is worth mentioning that 'pre-defined' should be at the design stage, not before the final analysis.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	710	711	4.1.3.2	The analysis and reporting should confirm a dose-exposure-response relationship that establishes the effective dose(s).' Why in case of biomarker it is necessary to confirm the established DER relationship, while for clinical endpoint, it is not necessary to establish DER relationship (line 536). Not clear. It should be explained. Moreover, this statement appears inconsistent with previous statements (e.g., lines 534-536) in the guideline that explicitly note that it may not be always feasible or even ethical to establish E-R relationships in pediatric clinical investigation. Finally, the analysis should confirm the dose-exposure-response relationship, but on p. 28 – 29 it is explained that in certain cases only PK data are needed if there are good reasons to assume that the same exposure results in the same efficacy. This seems contradictory with the requirement to confirm a dose-exposure-response relationship. Consider removing the statement or amending it.	The analysis and reporting should confirm a dose-exposure-response relationship that establishes the effective dose(s). OR: The analysis and reporting should confirm a dose-exposure-response the modeling relationship in the reference population that establishes the effective dose(s) in children.
GPT	713	723		Again, the differentiated approach is supported.	We propose to consider single-arm-studies as a potential stand-alone approach, e.g. in cases of known or proven good tolerability.
EFPIA	719	721	4.1.4	There can also 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 (see section 4.3.1).' A controlled pediatric study may evaluate efficacy and PK simultaneously. Also, a pediatric study may evaluate efficacy and PK in the different stages of the study. hence our proposed revision.	There can also 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 (see section 4.3.1). A controlled pediatric study may evaluate efficacy and PK simultaneously. Also, a pediatric study may evaluate efficacy and PK in the different stages of the study.
Koop Phyto	721	721	4.1.4	The PK/PD based approach is not suitable e.g. for locally applied medicines with local effects, vaccines, well established drugs or herbal medicinal products. Other tools such as empirical approaches need to be used and accepted. As a retrospective approach, the systematic collection of the existing experience from previous therapeutic use in children in a scientific manner can provide valid information on the safety and the therapeutic usefulness of such products.	Please add: Also other approaches including pro- and retrospective non-interventional studies (NIS) and aggregated information on medical experience can be part of the extrapolation plan (see section 4.3.)
Lundbeck	721	723	4.1.4	Given that the extrapolation concept and plan are likely to be updated regularly as new information comes in, would it be an option to outline several pediatric development scenarios that can be followed depending on outcomes of incoming data or other studies in the extrapolation plan? A pre-specified decision tree that is agreed upon with the authorities could help to select between development options efficiently.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
German Pharmaceutical Industry Association	725	774	4	The differentiated approach is strongly supported.	
AESGP	725	774	4	The differentiated approach is strongly supported.	
EFPIA	725	775	4,2	The 4.2 section should not be stand-alone but integrated in both the extrapolation concept and extrapolation plan sections. This section is clearly one of the most important ones in an extrapolation guidance. Unfortunately, the content is really vague (to put it politely) and does not provide any guidance at all. Conceptually, this section should explain (1) what models can be used for, (2) the need to validate the models, (3) the need to qualify the models for their use, and (4) key aspects which the models need to address. (1) In the context of pediatric development plans, the main use for models is to estimate the initial pediatric dose, and to estimate or predict a pediatric effect of the drug. Since the models will usually be built on data from the reference population, estimation of an initial pediatric dose or estimation of the drug effect in the pediatric population require extrapolation from reference to target population. (2) Validation of models can be done via standard approaches, for example by splitting the reference data into a training and a validation data set. Once validation is achieved, we know that the models describe the adult data well. (3) Since one is going to use the models for extrapolation to the target pediatric population, one needs to qualify the models for this purpose. In case of the estimation of the initial pediatric dose, this qualification of the model can be achieved by conducting a pediatric PK study, and by demonstrating that the observed exposures from this study indeed meet the objectives. It appears that this section describes certain aspect of PK studies, therefore may not be appropriate to be listed in parallel with "dose selection" and "efficacy studies"	It is suggested merging Section 4.1.3.1, 4.1.3.2, 4.1.3.2, and Section 4.2 into a new Section 4.2 titled "PK Studies', where "Model-Informed Approaches' can be a subsection
EFPIA	731	732	4,2	Modeling and simulation can be used to validate the pediatric extrapolation concept after completion of the pediatric study.' What does this sentence mean? Please clarify.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	733	735	4,2	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 and the simulation set up are clearly reported.' Fully agreed. And I would have hoped for guidance on how to do that, beyond just a statement that you have to do that.	
EFPIA	735	737	4,2	Typically, this information would be provided in the form of a modeling and simulation plan that the sponsor generates for internal documentation purposes but is also suitable for interaction with regulators.' Mention also simulation report, not only simulation plan.	Typically, this information would be provided in the form of a modeling and simulation plan and simulation report, if available, that the sponsor generates for internal documentation purposes but is also suitable for interaction with regulators.
EFPIA	746	748	4,2	"When using existing models (population PK, PBPK, population PK/PD models, etc), the specific characteristics of the target population, such as relevant body size and organ maturation, should be incorporated in the model." Consider adding proposed sentences to the section that discusses application of PBPK models.	PBPK models should be informed by a knowledge of ADME mechanisms and associated ontogeny of relevant drug-metabolizing enzymes and transport proteins. To increase confidence in pediatric PK predictions with PBPK models, they should be qualified for their ability to predict PK in adults and any available PK data in older pediatric subsets (e.g., adolescents).
EFPIA	750	752	4,2	For example, using models based on the reference population, analysis with pooled datasets, or Bayesian approaches with prior distributions for model parameters.' The text refer to techniques that can be used to incorporate information from the reference population in the analysis of the target population	Propose including some references relating to these techniques to help guide the reader.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	759	761	4,2	Not all data and model elements are equally valuable; therefore, assumption testing is an important aspect of any extrapolation exercise and should be integrated into the analysis plan and report.' We agree that assumption testing is an important aspect. however, we don't understand how this relates to the first part of the sentence "not all data and model elements are equally valuable". And since assumption testing is an important aspect, maybe spend a few sentences on how to do this, rather than just stating that one needs to do this.	
EFPIA	765	768	4,2	This section could be clarified further. It is important to distinguish between variability (between-subject) that naturally exists in the population and modelling or measurement error uncertainty that is model or instrument specific.	
EFPIA	771	771	4,2	"Where there is more limited or no data to support values chosen". Please consider modifying the sentence as proposed.	Where there is more limited or no data to support values chosen

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	772	774	4,2	"All of these can contribute to overall uncertainty in the results, and the different contributions that these could have should be addressed and justified during the exercise." Please consider modifying the sentence as proposed.	All of these can contribute to overall uncertainty in the results, and the different contributions that these could have should be addressed, and justified during the exercise and accounted for in model-based simulations as much as possible.
ProPharma Group	776	831	4,3	Can examples be provided where each efficacy study design would be considered appropriate. This will help drug developers to understand whether the extrapolation concept/plan is likely to be acceptable.	n/a
EFPIA	776	964	4,3	It appears that this section (Efficacy Studies) describes certain aspect of PK studies, therefore may not be appropriate to be listed in parallel with "dose selection" and "efficacy studies". It may be worth though mentioning that "The less data support is for quantifying the uncertainties, the more sensitivity analysis is needed." The thinking of this section is too much driven by the concept of having to compare the experimental treatment in pediatric patients with a control (either via a threshold, using external control data, or using a concurrent control). However, under an extrapolation concept one wants to extrapolate efficacy from a source (usually adult) to the target (pediatric) population. Therefore, in our view the primary objective of the analysis of the pediatric study data should be to provide evidence for similarity of efficacy. There are some important points with this regard mentioned in this section (lines 868-876), but in our view the entire section should focus on exactly this. With this in mind, the entire discussion on sample size that pops up here and there in this section should be revised, and whilst doing this, it should be kept in mind that there are often practical issues that make it impossible to recruit the numbers that result from any sample size calculation, even if type I error is elevated to 20% (say) or if non-inferiority margins are relaxed. (See comment below which is also part of this one)	Suggest merging Section 4.1.3.1, 4.1.3.2, 4.1.3.2, and Section 4.2 into a new Section 4.2 titled "PK Studies', where "Model-Informed Approaches' can be a subsection. Also, we would recommend to adapt the structure of section 4.3 as followed: 4.3.1 - Trials design (being either Uncontrolled Efficacy Studies, Externally Controlled Studies or Concurrent Controlled Efficacy Studies) ==> For each of the options, highlight powering to support decision making can be made with the specificities 4.3.2 - Frequentist or Bayesian assumptions and decision criteria ==> This section should be: i. to express that both approaches are appropriate to use depending on the design and available data. In the case of Frequentist, Type I error needs to be justified and agreed likewise for Bayesian, level of borrowing needs to be justified; ii. the need to pre-specify decision criteria and evaluate operational characteristics of the proposed design (false negative, false positive) iii. incorporation of external data that can use either a Bayesian or Frequentist approaches 4.3.3 - Analysis, Reporting, and Interpretation

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	776	964	4,3	With respect to study design, the only aspect that is covered is the choice of control group (i.e. single arm study analyses against a threshold, single arm study compared to external control data, and study with concurrent control). But there are more design considerations that need to be covered. For example, the option of unequal randomization (i.e. less patients to the control and borrowing of historical control information. More importantly, in many pediatric efficacy studies PK will play a key role, because the study will also be used to confirm that the initial pediatric dose that was selected based on modeling and simulation indeed meets the target (i.e. matches the exposure observed in adults). If so, there should be an early interim analysis. After all, we don't want to learn that the initial pediatric dose does not meet the target at the end of the efficacy trial – we want to learn about this as early as possible to be able to adjust the dose (which is also an ethical question to me). And there should also be an adaptive component to such trials – if the interim analysis does not confirm that the initial pediatric dose matches adult exposure, a revised dose should be proposed (using M&S and a revised PK model) and the study should be continued with this dose. Consider specificities linked to each statistical approach used and adapt the structure accordingly. It seems that not enough statistical methods are described in this whole section. For example, Bayesian methods such as power prior or meta-analytic predictive prior and hybrid methods such as propensity score stratified power prior are not mentioned. It is thus suggested providing more methodologies that can be applied in efficacy studies.	
EFPIA	777	847	4.3, 4.3.1, 4.3.2,4.3.3,4.3.4	All these section/subsections discuss an important design element which is the choice of control arm, but the current layout seems to narrow the whole section 4.3 on this element, which should be improved. Selection of endpoint is also one of the most important design decisions, therefore it would be helpful to include another section on this topic. For example, what if an endpoint is appropriate as the primary endpoint in the target population but not the reference population? Any guidance on such question would be extremely helpful.	It is suggested merging into a new Section 4.3.1 titled 'Choice of Control Arm' and each section can be subsequent subsection. Also Section 4.3.4 should be before the other design options to improve the flow.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	778	780	4,3	The options may include a randomised concurrent control, a formal statistical comparison against an external control, or a single arm trial.' One other option for efficacy studies is a randomised placebo-controlled study. It could be still ethical to have a placebo in a paediatric study for a few disease. For instance, we had performed a phase III placebo controlled pediatric study in depression, all patients received a psychosocial counseling in parallel to the study treatment. Theoretically, a comparison against concurrent/non-concurrent control from a platform trial may be possible. Consider adding platform trials to compare against concurrent/non-concurrent control.	….a randomised concurrent control or placebo control, a formal statistical comparison against an external control, or a single arm trial.
GPT	783	793		The single-arm-study-approach is appreciated. Safety considerations are, however, missing. Long-standing evidence of safety may suffice to justify the use of a medicinal product in the target population.	We suggest to consider adding long-standing safety evidence as additional criterion for the use of a single-arm-study-approach. A risk-based approach should be used when determining the extent of additional data.
German Pharmaceutical Industry Association	783	793	4	The single-arm-study-approach is appreciated.	
AESGP	783	793	4	The single-arm-study-approach is appreciated.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	784	786	4.3.1	We agree that single arm studies may be the most appropriate design when running a pediatric efficacy trial under an extrapolation approach. However, we think the justification (standard of evidence in the reference population being a single arm trial) does not make sense. There are other more important reasons why to conduct a single arm trial. For example, a registered control for children may not exist, and a placebo control may be unethical. As another example, lack of ability to recruit may be a reason for running single arm trials in pediatrics. This may even be the case in non-rare pediatric indications, because there may be too many pediatric studies ongoing in one indication at the same time, or because there is already a registered good treatment option and parents may be reluctant to consent to a clinical trial. Is this the only scenario that a single arm efficacy study is appropriate?	
Agios	786	787	4.3.1	Suggest rewording for clarity as in some cases a "threshold" may not be used but rather a criteria should be pre specified and justified.	When designing the study, the criteria to determine whether or not the primary objective was met should be pre-specified.
EFPIA	786	787	4.3.1	When designing the study, how the primary efficacy objective would be evaluated should be defined using a pre-specified threshold.' Meaning is unclear. Is the proposal of the study objective to demonstrate that the treatment effect is above/below a pre-defined threshold? The description of the objective and sample size justification described on Line 789 is not consistent. Consider repharsing.	When designing the study, how the primary efficacy objective would be evaluated should be defined, e.g. mean treatment effect is above/below using a pre-specified threshold or level of precision.
Agios	789	790	4.3.1	This concept is associated with the concept in line 787 and suggest it be combined into the same paragraph. Suggest rewording as show to link to the primary objective evaluation concept	The sample size of the studies should be calculated and justified to ensure the primary objective can be assessed (eg, estimates have adequate precision to enable a decision to be reached as to whether or not the primary objective was met)

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	789	792	4.3.1	The sample size of studies should be calculated to ensure the threshold is met, or to ensure that an estimate of sufficient precision is obtained.' This discussion of threshold is unclear. How will the threshold be established or justified? What does a sufficient precision mean? Give a little more guidance for what is expected regarding precision. Sample size calculation for a single arm trial can ensure sufficient precision but can't ensure that the threshold is met.	The sample size of studies should be calculated to ensure the threshold is met, or to ensure that an estimate of sufficient precision is obtained, so that there is a high probability that the pre-specified threshold is met under certain assumptions.
EFPIA	790	793	4.3.1	Clarification proposed with regards to the use of RWD.	External data can be used to contextualise the results (e.g., using published literature to understand the context of the results of the study with respect to current clinical practice, but without requiring a formal comparison of efficacy to external data) or available adequate real-world data sources (e.g., electronic medical record, claims database or registries).
GPT	795	801		Externally controlled studies may, among others, comprise real world evidence sources. However, to our understanding real world evidence can also consist e.g. in NIS, registries, or prescription data.	We recommend including the mentioned examples for real world evidence as well.
EFPIA	795	964	4.3.2- 4.3.7	Although use of external data as a comparator is discussed in multiple sub-sections, there is not much guidance on methodologies or examples for such analyses. Demand for external data for pediatric extrapolation is high and some guidance on methodologies would be appreciated. Consider adding references or examples of methodologies to analyse efficacy data in children with external data as formal comparator.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	795	795	4.3.2	It would be helpful to elaborate more on the difference between externally controlled and single arm studies. Both type of studies assign all enrolled subjects to the treatment arm. The difference is in the analysis. In single arm efficacy studies, external data can be used to define the threshold on efficacy endpoint. Aggregate data can be used. In externally controlled studies, individual patient level external data are selected and used as controls in the analysis. Do we use externally controlled study design to allow us to choose control subjects that match the baseline characteristic of the subjects enrolled in the study?	
EFPIA	796	797	4.3.2	It may be possible and appropriate in some circumstances to use external data as the formal comparator in a trial.' By definition the external data would not be "in" a trial. Change the word "in" to "for".	It may be possible and appropriate in some circumstances to use external data as the formal comparator in for a trial
EFPIA	798	798	4.3.2	"This could be from the comparator arm in the reference population, relevant control arms from other randomized controlled trials (RCTs), or real-world evidence sources in the target population." Unclear to which population the following sentence refers to: "relevant control arms from other randomized controlled trials (RCTs)"; historical control arms from RCTs within the same disease? Consider amending the sentence. In addition, consider highlighting that prospective observational real world evidence studies often offer significant advantages in terms of collecting data beyond standard of care, compared to retrospective RWE studies. See proposed addition.	This could be from the comparator arm in the reference population, relevant historical control arms from other randomized controlled trials (RCTs) in the target population, or real-world evidence sources in the target population. Prospective real world evidence studies can offer advantages over a retrospective external comparator in collecting data beyond standard of care.
Lundbeck	799	801	4.3.2	This seems to suggest that only "use of external data beyond these sources" needs to be justified, should one not always provide a justification when using any kind of data?	
EFPIA	804	806	4.3.2	Since the data are compared directly with a data source external to the study, appropriate statistical methods should be used to account for differences between the populations.' Examples of appropriate statistical methods should be provided.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	806	807	4.3.2	It is important to reflect that these studies would still be controlled, albeit with a non-randomized control, which differs from the approach of just comparing to a threshold.' The sentence is somewhat unclear/imprecise. consider revisiting.	It is important to reflect that these studies would still be controlled, albeit with a non-randomized control, which differs from the approach of just comparing to a concurrent control where the treatment effect can simply be compared against a respective threshold.
German Pharmaceutical Industry Association	809	831	4	Guidance for Concurrent Controlled Efficacy Studies is very helpful.	
AESGP	809	831	4	Guidance for Concurrent Controlled Efficacy Studies is very helpful.	
EFPIA	809	818	4.3.3	This section considers demonstrating success in a powered study. In some situations, where sample size is limited, powering the study might not be feasible. Then justification of sample size might aim at achiving a pre-specified precision in determining the treatment effect.	Add possibility to justify sample size by considering the precision in determining the treatment effect for investigational treatment vs. control.
EFPIA	809	809	4.3.3	It would be a good introduction to section 4.3.3, if in the sequel you would be discussing the situations were randomized controlled efficacy studies would be needed, if you would explain how the study designs might be different than those required in the reference population, and if you would elaborate on what you mean by the last two sentences. Since examples using frequentist terminology are used, it should be mentioned that Bayesian approaches with weak priors, i.e. where inference is determined by data, could be used as well to demonstrate success. There is no clear guidance whether a strict control of the type I error is required. Consider adding the proposed sentence.	Bayesian approachs with weak prior (inference is mainly determined by data) can be used to demonstrate treatment success as well. For this a minimum posterior probility of 95% is recommended to demonstrate success. Lower values should be justified.. A strict control of the type I error is not required and an inflation of the type I error is accepted to a certain extend which has to be discussed with regulatory agencies."

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	812	814	4.3.3	Based on the pediatric extrapolation concept, the need for controlled studies and the ability to extrapolate leads to study designs different than those that were required in the reference population.' Is the study design always different than those that were required for the reference population?	Clarify whether similar study desings in target population and reference populations are aceptable as well.
Lundbeck	814	818	4.3.3	This seems sensible, but it is very difficult to choose what level of false positive and negative rates are then to be used. Some guidance on how to do this or what kind of arguments could be used would be nice.	
EFPIA	814	816	4.3.3	This will lead to a different relationship between the false positive rate, the false negative rate and sample size that is not the same as it is in the reference population.' This part of the section is unclear. Is this section restricted to dichotomous endpoints, what are false positive, false negative in the context of concurrent controls?	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Agios	816	818	4.3.3	Could be helpful to provide an example as to what is meant since in a frequentist setting we use Type I and Type II error but in a Bayesian setting other concepts may be used.	When the sample size is limited, the relative importance of false positive and false negative results should be considered carefully (eg, Type I error and Type II error in a frequentist approach).
EFPIA	822	823	4.3.3	The extrapolation approach will result in a sample size smaller than one would expect for a standalone efficacy study.' This may be true sometimes but may not always be true. I would imagine that whether there is a sample size reduction depends on the knowledge gap, uncertainties in the extrapolation, and the study design being considered for the standalone study. Also, if the extrapolation assumptions do not hold, then there will be a bias and will lead to more studies/additional sample size. Is this because there is additional information to be borrowed from the adult study? See suggested revision.	The extrapolation approach may reduce the will result in a sample size needed to fill the knowledge gap relative tto using smaller than one would expect for a standalone efficacy study.
EFPIA	823	825	4.3.3	If the study is powered to meet a relaxed success criterion with a significance threshold larger than 0.05, this should be justified in advance.' The guideline expects an upfront agreement in case of a frequentist design when a different threshold than p=0.05 is used while for a Bayesian design extend of borrowing and interpretation is pre-specified only. A similar wording might be used for both approaches (frequentist and bayesian). Also, quantifying the risk of false postitive by a significance level requires a more detailed specification of the test hypothesis, i.e. one-sided ot two-sided test. See proposed revision.	If the study is powered to meet a relaxed success criterion with a significance threshold larger than 0.05 for a one-sided test, this should be justified in advance.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	827	830	4.3.3	An alternative approach for active controlled trials may be to maintain the conventional type I error rate but widen the non-inferiority 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.' Please clarify, what does “widen the non-inferiority margin” and "consistensy" mean in this context? For clarity, the sentence could be revisited as proposed.	For aAn alternative approach for actively controlled non-inferiority trials may be to maintain the conventional type I error rate but widen the non-inferiority a wider margin usually used in than the de novo adult development, or a smaller confidence interval (e.g, 90% CI instead of 95% CI) can be used, 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.
EFPIA	830	831	4.3.3	'It will be important to ensure the point estimate obtained should be consistent with that in the reference population.' It is not entirely clear what is proposed in this paragraph. Do you mean that instead of showing superiority it can be sufficient to show non-inferiority, and then to require that the point estimate for efficacy is in favor of the experimental treatment arm as it was for the reference population? Or does it only refer to situations where the adult trial already used a non-inferiority design? Please clarify when non-inferiority design can be used and the meaning of consistent point estimate.	It will be important to ensure the point estimate obtained in the target population should be consistent with that in the reference population
Koop Phyto	832	832	4.3.4	Rationale see line above	Please add new chapter 4.3.4: **Subsequent NIS-based iterative age extrapolation of clinical evidence** The collection of product-specific PK/PD data for use as surrogate parameters as the only tool for determining a
EFPIA	833	847	4.3.4	The guideline should give some examples of selection criteria for which data/studies to include.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	835	836	4.3.4	"When identifying which information will be incorporated into the analysis of the pediatric study, relevant data should be identified through a systematic search using pre-specified selection criteria." Systematic search applies more to publication or literature data, however, it should not be applied to all situations, e.g. the adult trial is appropriate and does not have to be identified through a systematic search. See proposed change.	When identifying which information will be incorporated into the analysis of the pediatric study, relevant data should be identified through an appropriate approach (e.g. a systematic search using pre-specified selection criteria for data from the literature)."
EFPIA	845	847	4.3.5	"Potential differences between the study from which the reference data will be derived and the data generated in the target population can be adjusted and accounted for in the analysis as much as possible" The reference to "adjustment" should be required rather than optional. Also, reference should be made to the use of appropriate methods, for clarity.	Potential differences between the study from which the reference data will be derived and the data generated in the target population should can be adjusted and accounted for in the analysis as much as possible, using appropriate methods.
EFPIA	848	848	4.3.4	It would be useful to add a paragraph about incorporating external data based on Bayesian approach. Consider addind the proposed statement.	Add this paragraph at the end of Section 4.3.4, line 848: When the degree of similarity between external data and the target paediatric population is in a continuum, Bayesian approaches such as Bayesian hierarchical models and dynamic borrowing using commensurate priors, for example, can be used to integrate and
ProPharma Group	849	857	4.3.5	Wording currently states: "If the available information (based on reference data, or outputs from a modeling and simulation exercise) is summarised as a statistical distribution then the effective sample size is a good way of describing how much information is being used." It would be helpful if further quantification could be given. For drugs being developed for orphan diseases, often the data available are minimal as the clinical development program is small.	n/a
EFPIA	849	849	4.3.4	Incorporating external data or augmenting control arm by external data might lead to an inflation of the type I error. It should be clarified whether a strict control of the type I error is required. Consider addind the proposed statement.	Add "A strict control of the type I error is not required and an inflation of the type I error is accepted to a certain extend which has to be discussed with regulatory agencies."
EFPIA	850	857	4.3.5	We guess that this paragraph refers to a specific situation, where Bayesian analysis with informative priors is being used? Anyway, can you provide the context in which this paragraph applies. I am not sure that the paragraph makes sense in other contexts. For example, in the context when the primary objective of the analysis is to demonstrate similarity between reference and target population outcomes, I don’t see how this paragraph would apply. (In such a context one could use the adult data to build an exposure-response model (including additional baseline risk factors as covariates) and use this model to predict the outcome of a pediatric trial to compare observed with predicted outcomes.)	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	850	851	4.3.5	"It is important to understand a priori how much available information is being incorporated into the design and analysis to support the interpretation of the pediatric trial." Clarifying "a priori" to mean "prior to seeing the post-treatment efficacy data" would be helpful. As commented later in this paragraph, quantifying information such as effective sample size calculation would not be possible without seeing the baseline prognostic factor distribution from both groups. External data can be used in a single arm study to define threshold, as external controls, or to define priors for the outcome variables in the Bayesian analysis. Does this section refer to all of the above use? Are there any other use? It is suggested to change "a priori" to "prior to seeing the post-treatment efficacy data"	It is important to understand a priori prior to seeing the post-treatment efficacy data, how much available information is being incorporated into the design and analysis to support the interpretation of the pediatric trial
EFPIA	851	854	4.3.5	The meaning of the sentence is unclear, especially ".... but also how much of the data generated in the reference population is relative to the amount of data generated in the target population." A rewording would be welcome for more clarity.	
EFPIA	854	857	4.3.5	"If the available information (based on reference data, or outputs from a modeling and simulation exercise) is summarised as a statistical distribution then the effective sample size is a good way of describing how much information is being used." There are many different methods to summarise the effective sample size of a probability distribution which can lead to very different results for non-conjugate distributions. The tool may not be as useful as it might seem to be unless the method is harmonised. Is there a preference for a specific method?	Please clarify.
EFPIA	856	856	4.3.5	This section started with a statement that the amount of information should be understood a priori, however, the effective sample size can only be calculated after we make the comparison between the target and reference population's characteristics. Moreover, "statistical distribution" is not a well defined technical term.	If the available information (based on reference data, or outputs from a modeling and simulation exercise) is summarised as a statistical probability distribution then the effective sample size is a good way of describing how much information is being used.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	859	861	4.3.5	"If Bayesian approaches are used, different ways of using the prior information, for example by using a mixture prior or power prior, will have a different effective sample size depending on the choice of parameters used in the model." Consider revising the sentence as the impact on the sample size is not necessarily linked to the use of a Bayesian approach but to the fact of using a model based approach	If Bayesian approaches are a model-based approach is used, leveraging external data, different ways of using the prior information, for example by using a mixture prior or power prior, will have a different effective sample size might depending on the choice of parameters used in the model.
EFPIA	859	866	4.3.5	Effective Sample Size (ESS) is not defined for all prior structures, specifically mixture priors are called out here and there is no exact calculation for ESS under this methodology. Consider including other examples of how the reference data is impacting the analysis of the source data, such as measuring the influence on the posterior (point estimates, bias, error/precision) under various scenarios. ESS can be a nice metric for fixed-borrowing approaches but an extension of this concept is needed for dynamic borrowing approaches (where ESS is a function of how similar the likelihood and the prior are). Can you address this somewhere, so that there is not a perceived requirement to report ESS when using something like a mixture prior. Or perhaps clarify that approximations to ESS may be necessary depending on the methodology being considered.	
EFPIA	866	866	4.3.5	Clarification added; consider adding the proposed sentence.	Different statistical methods for effective sample size calculation might also yield different results.
EFPIA	878	884	4.3.5	It is possible to base a pediatric extrapolation plan using a biomarker, surrogate endpoint, or clinical endpoint as the primary endpoint in the target population even when the primary endpoint in the reference population differs. In this scenario, the guideline states that an evaluation of robustness of the correlation should be conducted, and sponsors may incorporate potential pediatric endpoints into the adult development program. It is unclear how robust the correlation will need to be in order to gain regulatory agency acceptance.	Include a general statement regarding how robust the correlation will need to be and what evidence is needed in order to gain regulatory authority acceptance. Furthermore, include guidance around incorporating potential pediatric endpoints in the adult development program. For example, to best support the pediatric extrapolation plan, pediatric endpoints in the adult program should be assessed in adult participants 18-25 years of age.
Agios	895	896	4.3.6	The statement "additionally the results of an analysis of the data alone should always be provided", would be more clear with examples.	Additionally the results of an analysis of the data alone (i.e., frequentist approach or Bayesian approach using non-informative priors only) should always be provided.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	895	896	4.3.6	"If a Bayesian design is used, the full operating characteristics should be provided. Additionally, the results of an analysis of the data alone should always be provided." It is not clear what full operating characteristics means. Consider to give more details or a reference. Does "results of an analysis ofthe data alone" refer to frequentist analysis or simple summary statistics? It seems redundant to analyse the data using both frequentist and Bayesian approach. Please clarify and consider proposed revisions.	If a Bayesian design is used, then should an analysis with non -informative or "flat" prior be provided or should the data be aggregated by mean, sd, etc, the full operating characteristics should be provided. Additionally, the results of an analysis of the data from the target population alone should always be provided.
German Pharmaceutical Industry Association	898	942	4	The guidance Analysis, Reporting, and Interpretation is highly appreciated.	
AESGP	898	942	4	The guidance Analysis, Reporting, and Interpretation is highly appreciated.	
EFPIA	898	964	4.3.7	This section separates efficacy study designs into frequentis and Bayesian. However, recently more and more designs are hybrid (e.g. design using propensity score stratified power prior approach), therefore it is worth mentioning this trend and describing considerations for such hybrid designs. Suggested references on hybrid designs: Baron, E., Zhu, J., Tang, R., & Chen, M. H. (2022). Bayesian divide-and-conquer propensity score based approaches for leveraging real world data in single arm clinical trials. Journal of Biopharmaceutical Statistics, 1-15. Liu, M., Bunn, V., Hupf, B., Lin, J., and Lin, J. 2021. Propensity-score-based meta-analytic predictive prior for incorporating real-world and historical data. Statistics in Medicine, 40(22): 4794-4808 Wang, C., Li, H., Chen, W.-C., Lu, N., Tiwari, R., Xu, Y., and Yue, L. Q. 2019. Propensity score-integrated power prior approach for incorporating real-world evidence in single-arm clinical studies. Journal of biopharmaceutical statistics, 29(5):731–748.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	898	898	4.3.7	Suggest to add a statement on the role of estimands, and any special considerations that would be needed, when developing estimands in the context of extrapolation especially when leveraging external data or applying Bayesian methods.	
Takeda	899	902	4.3.7	Agreement with whom is unclear. Health authorities?	
EFPIA	899	902	4.3.7	"If a frequentist design is used, an alternative threshold to cross other than the standard two-sided significance level of 5%. should be agreed upon in advance and a frequentist analysis compared to this alternative threshold provides a justification of the pediatric extrapolation concept." We found this paragraph confusing. It was already mentioned earlier that "If the study is powered to meet a relaxed success criterion with a significance threshold larger than 0.05, this should be justified in advance." it is suggested to have it reworded. Sentence is hard to understand and might be incomplete. Clarification of the text is recommended, as outlined.	If a frequentist design is used, an alternative threshold to cross other than the standard two- sided significance level of 5% should be agreed upon in advance and a justification of a frequentist analysis compared to this alternative threshold should be provides d in the a justification of the pediatric extrapolation concept." keep the comment but delete the proposed revision as not clear.
Agios	900	900	4.3.7	Typo, please delete the period after "5%"	
EFPIA	900	900	4.3.7	Typo: Delete '.' after 5%.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	903	905	4.3.7	"A frequentist meta-analysis approach combining reference and target data could be conducted if it is appropriate to formally analyze the data together" This sentence is an example for the limited perspective of this guidance. It assumes that one needs to compare the experimental treatment with a control and attempts to gain power by appropriately integrating external data from the reference population into the analysis. However, an extrapolation concept attempts to extrapolate efficacy from the reference population to the target population. This requires a model, which needs to be qualified for extrapolation. Therefore, a pediatric data set (which would usually be small as compared to the reference data set) should primarily be used to qualify the model for extrapolation. Formally analyzing data together via a meta-analysis would only be appropriate if the similarity of response to intervention between reference and target populations has been demonstrate. You should discuss how to assess this in the guidance. Same comment applies if Bayesian methods are used.	
EFPIA	903	905	4.3.7	"A frequentist meta-analysis approach combining reference and target data could be conducted if it is appropriate to formally analyze the data together" Meta-analysis combining adults and target data would need a strong rationale of similarity between both populations. Maybe "if it is appropriate" could be developped at least by "if population sufficiently similar. Moreover, clarification added with a proposed addition.	A frequentist meta-analysis approach combining reference and target data could be conducted if it is appropriate (at least populations with sufficient similarity) to formally analyze the data together. Pharmacometric methods such as pooled population PK/PD modeling on patient-level datasets, or model-based meta-analyses on summary trial-level data may be considered.
Agios	908	909	4.3.7	To avoid misunderstanding, should be clarified that the analysis can be updated as external data are generated (not after the data from the pediatric trial is available)	This analysis should be pre-specified and may be updated as additional external data are generated or become available.
EFPIA	908	909	4.3.7	"This analysis should be pre-specified and updated as data are generated." Does "updated as data are generated" mean the analysis plan should be updated as new external data are generated? The analysis plan shouldn't be updated as new trial data are generated. This sentence should be deleted or clarified.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	910	912	4.3.7	"Plots of posterior distributions resulting from Bayesian analyses may better contextualize the summary statistics derived from Bayesian distributions." Consider rephrasing as suggested.	Plots of posterior distributions resulting from Bayesian analyses may better contextualize the summary statistics derived from resulting from Bayesian analyses distributions.
EFPIA	912	915	4.3.7	"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 and where they were reported, along with a justification as to why they are considered to be appropriate to include."	
Lundbeck	919	924	4.3.7	Could you say something about the ”pre-specified amount” for the Bayesian approach? Many Bayesian approaches will calculate the probability that the treatment is better, corresponding to zero as the pre-specified amount. The comment on non-inferiority indeed corresponds to a negative pre-specified amount	
EFPIA	926	927	4.3.7	"It is important to understand how similar the target data are to the reference data and to use metrics to define such similarity." Does "results of an analysis of the data alone" refer to frequentist analysis or simple summary statistics? It seems redundant to analyse the data using both frequentist and Bayesian approach. An example would be helpful.	It is important to understand how similar the target data are to the reference data and to use metrics (e.g., xxxx) to define such similarity.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	927	929	4.3.7	"If the observed data in the study are not similar to 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." 'data... are' reads strange. Consider revising as proposed.	If the observed data in the study are is not similar to the observed reference data, this may…
Lundbeck	931	935	4.3.7	It seems a bit odd to justify extrapolation, but then to have to do this after all. Could you elaborate a bit on why this would be of interest?	
Agios	932	935	4.3.7	"but statistical significance without borrowing has failed to be achieved due to a small sample size". The "statistical significance" implies frequentist approach, but the "borrrowing" and "weight" implies Bayesian approach. Change "statistical significance" to "success criteria"	Nevertheless, if the data in the target population is 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).
EFPIA	935	935	4.3.7	A "tipping point" is the minimum shift to overturn the statistical significance (i.e., a significant result would become a nonsignificant one), but I believe the described context in this paragraph is the opposite (i.e., seeing how much better it needs to be for a non-significant result to become significant). Please clarify.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	945	945	4.3.7	"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, and type I error rate control." For the informative prior, it is quite usual to question the choice of the prior distribution. However, besides that, it is also important to know how the source data are selected. To keep the integrity of the analysis, the source data need to be representative. consider adding our proposed statement.	Justification of the choice of the source data needs to be provided. 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, and type I error rate control.
EFPIA	945	947	4.3.7	"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, and type I error rate control." It is suggested that guidance be included regarding the use of clinical trial simulations. Such analyses should be carefully selected to investigate the assumptions made with the primary estimator and other limitations with the data.	Sufficient clinical trial simulations should be conducted before or during protocol development, by considering optimisation of study design, amount of data which is reasonable to borrow from the reference population, sample size and power, endpoint selection, appropriate Bayesian success criteria, comparison with frequentist method if applicable etc.
EFPIA	948	951	4.3.7	"As an example, one possible method amongst many is to use a robust prior: a two-component mixture prior where one component is an informative prior based on the source data and the second is a weakly informative prior independent of the source evidence." Do robust mixtures always need to be 2-component mixtures? So long as there is a weakly informative component being mixed with source data, does the source data always have to be a single component of the mixture or could it be multiple components (for example 2 informative components representing 2 reference trials mixed with 1 weakly informative component)? Much simulation would be necessary to justify the weights assigned to each component of a mixture, but there's no reason why it would have to be restricted to only 2 components (mathematically speaking). Consider our proposed revision. Does "a weakly informative prior independent of the source evidence" refer to a somewhat flat prior covering the plausible range of the parameter?	As an example, oOne possible method amongst many is to use a robust mixture prior: for example a two-component mixture prior where one component is an informative prior based on the source data and the second is a weakly informative prior independent of the source evidence.
EFPIA	951	952	4.3.7	"The weakly informative component should be carefully chosen to ensure adequate borrowing behavior." What does "to ensure adequate borrowing behavior" mean? The weakly informative prior is not based on source data, so it should correspond to a low effective sample size. How do we ensure adequate borrowing behavior? Please clarify.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Agios	958	964	4.3.7	The current paragrah starts with sensitivity analysis, but the second and third sentence are still referring to how a prior should be constructed and should be part of the paragraph that starts on Line 945.	Suggest moving the second and third sentence of the paragraph 958 to 964 to the end of the paragraph in line 956.
EFPIA	958	960	4.3.7	"A sensitivity analysis such as a tipping point analysis can be a useful tool for retrospectively assessing the robustness of conclusions to the strength of prior assumptions about similarity of source and target population parameters." Is the tipping point analysis to assess the robustness of the conclusions w.r.t the amount of borrowing? Which parameter should the tipping point analysis be done on? Please clarify.	
European Association of Hospital Pharmacists (EAHP)	966	1017	5	EAHP would like to offer following topics for the further consideration of EMA for addition to the very last part (5.; most of them are already elsewhere in the text, but we feel they should be more emphasised): While the text on page 34 (line 746 and following) mentions organ maturation and body size, I felt this should be more emphasised as a certain limitation to the extrapolation (or that more cautious approach should be taken), especially with concern to premature newborns. Another thing to consider is that bioavailability of the medicines can differ between adult and pediatric populations (plus the dosage form plays very important role here; and very often there is a lack of the dosage form intended for children). Another point that should be more emphasised in the text is the need of promoting/education of the carers in relation to the pharmacovigilance plans with necessary emphasis on possible delayed or long-term toxicity/adverse events.	
EFPIA	974	975	5,2	"Even when extrapolation of safety data is justified, there may be additional safety issues that should be addressed." We do not extrapolate safety data but rather extrapolate evidence or outcomes. As mentioned on lines 47, 348, and elsewhere, data itself is typically not extrapolated, but findings are. Suggest replacing with "Extrapolation of safety findings" or "Extrapolation of safety outcomes".	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	975	977	5,1	"A comprehensive safety plan, including the need for pre- and post-marketing safety data collection should be described in the extrapolation plan." What is meant by safety plan - is this a document that is similar to the EU risk management plan or relevant sections hereof? With an extrapolation plan created during early clinical development, the need for post-marketing safety data collection will most likely not be evident as potential risks for the product may not be known due to limited available clinical and nonclinical data	Please provide further guidance or more specific examples of what is meant by a comprehensive safety plan.
ProPharma Group	979	1017	5,2	It would be helpful to include in this section some indication of the anticipated timings. It is assumed that the Agencies would find it preferable for drug developers to submit and agree the paediatric plan prior to including adolescents in an adult trial and therefore as part of this, get agreement on the extrapolation concept as stated in line 993. However, as per previous comments and following experience with discussing extrapolation concepts/plans with the Agency, often at these earlier stages of development, the Agency are less inclined to agree to extrapolation until further adult data are available.	n/a
EFPIA	991	993	5,2	"The decision to include a pediatric cohort (e.g., an adolescent subgroup 12 to 17 years of age) in an adult (e.g., > 18 years of age) clinical trial assumes the disease and response to treatment are sufficiently similar between the adolescent and adult patients." 18 is excluded in the parenthetical text defining adults; change the inequality to include 18	The decision to include a pediatric cohort (e.g., an adolescent subgroup 12 to 17 years of age) in an adult (e.g., > ≥18 years of age) clinical trial assumes
EFPIA	999	1010		It is unclear what safety endpoints should be considered for extrapolation. Treatment Emergent Adviser Events (TEAE) is a common safety endpoint but could contain different AEs for target and source populations.	It is suggested including additional considerations on safety endpoints used for extrapolation
EFPIA	999	1004	5,2	"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 ICH E9)."	We suggest including additional considerations regarding adolescent sample size especially when included in adults studies as well as inclusion of adolescents in long term safety follow-up.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	1006	1017	5,2	To achieve getting data ealy with pediatric plans, sponsors and regulators should consider better timing of pediatric trials. For this reason we propose an additional sentence, reminding that early pediatric development plans would be helpful in this regard.	It is recommended that the sponsor engage early with Health Authorities to align on pediatric development plans that include adolescents in adult clinical trials.
EFPIA	1011	1012	5,2	" (6) the willingness of pediatric investigators to participate in a subsequent pediatric only trial that would now exclude adolescents." Suggest changing the word "would" with "may" since the sentence, as currently stated, implies that inclusion of adolescents in an adult trials precludes their inclusion in pediatric trials, which is not necessarily the case.	(6) the willingness of pediatric investigators to participate in a subsequent pediatric only trial that would may now exclude adolescents.
EFPIA	1012	1012	5,2	Also address potentially splitting the analysis of the pediatric population, making it more difficult to achieve statistical significance (or draw conclusions) on any single age subgroup.	"[…] and (7) splitting the statistical analyses between adolescents and younger children."
Koop Phyto			4.3.2	RWD include by far more types of data as compared to externally controlled studies as listed in 4.3.2, to which the link on RWD in 3.4 refers. Real-world evidence is the clinical evidence regarding the usage and potential benefits or risks of a medical product derived from analysis of RWD. RWE as defined by FDA can be generated by different study designs or analyses, including, among others, also observational studies (prospective and/or retrospective).	See comment on 4.1.4