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

11 February 2026
EMA/389217/2025

Overview of comments received on ICH E20 Guideline on adaptive designs for clinical trials (EMA/CHMP/ICH/206586/2025)

Please note that comments will be sent to the ICH E20 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
ACRO	0	0	0	Founded in 2001, the Association of Clinical Research Organizations (ACRO) is a non-profit trade association representing the world's leading clinical research and technology organizations, which provide specialized services that are integral to the development of drugs, biologics and medical devices that enable patients to live longer, healthier, and more productive lives. ACRO members provide a wide range of services and digital technologies across the entire spectrum of development - from pre-clinical, proof of concept, and first in human studies through post-approval, pharmacovigilance, and health data research. ACRO member companies employ nearly 400,000 people worldwide and conduct research in every global region.	
ACRO	0	0	0	ACRO supports the goal of establishing harmonized principles that encourage innovation while maintaining scientific rigor and trial integrity. Industry has long adopted adaptive designs but the absence of an international guideline setting the regulatory framework and main acceptance criteria of adaptive designs has been a hurdle for innovative clinical research. This E20 guideline provides clarity and will enable discussions about planning for adaptive designs. We appreciate that the guideline provides assurance that planning, intentionality, and mid-stream mitigating effort is relevant, as this will help to embed novel design considerations throughout the stages of clinical development, trial conduct, and the product lifecycle.	
ACRO	0	0	0	ACRO recommends that the final E20 guideline encourage innovation in trial design and execution while maintaining methodological integrity and patient safety. Adaptive designs, when appropriately pre-specified, can improve efficiency, ethical balance, and decision-making without compromising statistical validity. ACRO recommends providing additional clarity in the final guideline to ensure both scientific rigor and operational feasibility. In particular, the final guideline would benefit from: A clear definition of “interim analysis” Clearer expectations for digital system validation, version control and auditability, consistent with modern data governance standards. Explicit recognition of advanced adaptive methods, including platform, basket, and biomarker-driven trial models increasingly used in complex research programs. Further illustrative examples to support understanding and expectation, per subsection Explicit cross-referencing to related guidelines including ICH E6(R3), E8(R1), E9, M11, and M13 to promote consistent terminology, system validation expectations, and documentation standards across the development lifecycle.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
ACRO	0	0	0	The documentation requirements outlined in E20 are thorough but risk creating redundancy across regions. We recommend harmonizing expectations with ICH M11 and regional data submission templates, ensuring that required documentation reflects structured, digital formats rather than duplicative narrative reports – similar to the restructuring in ICH's M4Q(R2). This would strengthen transparency, support automation, and align with ICH's broader modernization strategy.	
ACRO	0	0	0	ACRO thanks the EMA for the opportunity to provide this feedback on ICH E20. Please do not hesitate to contact ACRO if we can answer any questions. Respectfully submitted, Karen Noonan, Senior Vice President, Global Reglatory Policy, ACRO	
INCiPiT (c4c-S)	0	0	0	Contextual and general comments: - potential increase in indirect costs: although overall time and costs are reduced, the complexity of adaptations may require greater investment in technology, training, and ongoing monitoring; - specific use for certain types of studies: adaptive designs are particularly suitable for studies with high uncertainty or in the early stages, but are not always appropriate for all trials, such as those with very rigid endpoints or rare conditions; - evolution of digital technologies: the increased use of digital systems for data collection and analysis facilitates the implementation of adaptive studies, but at the same time requires special attention to security and privacy; - cost implications: addressing how adaptive designs may impact trial costs and timelines, and how these factors should be considered in economic evaluations. - early and ongoing dialogue between trial sponsors and HTA bodies can enhance the relevance and quality of evidence; early engagement, initiating discussions with HTA bodies during the trial design phase to align on evidence requirements; regular updates: providing HTA bodies with interim results and updates on any adaptations to the trial design. - incorporating RWE can strengthen the applicability of trial results to broader patient populations; - integration strategies: methods for integrating RWE into adaptive trial designs. Regulatory and HTA acceptance: Clarifying how RWE is evaluated by regulatory authorities and HTA bodies. In summary, from an HTA perspective, the ICH E20 guideline should provide clear guidance on how adaptive designs can generate evidence that meets the rigorous standards required for reimbursement decisions. This includes ensuring robust clinical endpoints, considering economic factors, emphasizing transparency, encouraging early engagement with HTA bodies, and exploring the integration of real-world evidence.	
Breakthrough T1D	0	0		Breakthrough T1D (formerly JDRF) appreciates the opportunity to provide comments on the ICH E20 Guideline on adaptive designs for clinical trials, which has reached Step 2b of the consultation procedure. ABOUT BREAKTHROUGH T1D As the leading global type 1 diabetes (T1D) research and patient advocacy organization, Breakthrough T1D helps make everyday life with type 1 diabetes better while driving toward cures. We do this by investing in the most promising research, advocating for progress by working with governments to address issues that impact the T1D community, and helping educate and empower individuals facing this condition. Since 2015, our organization has invested more than €57 million in European projects. In addition, 30 clinical trials are currently funded by Breakthrough T1D in Europe. Breakthrough T1D welcomes the ICH initiative to develop internationally harmonized guidance on adaptive clinical trial designs. For studies involving small patient populations, such as single-arm, potentially curative T1D cell therapy trials, harmonized guidance across regulatory bodies on adaptive elements has the strong potential to accelerate regulatory review processes. This includes applications for clinical trials, scientific advice, special designations, and marketing authorizations, spanning the entire product lifecycle from early development through post-authorization requirements.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Breakthrough T1D	0	0		[Continued from Above] Adaptive trial designs are increasingly utilized in type 1 diabetes (T1D) research, particularly for disease-modifying therapies, due to their flexibility and efficiency. Adaptive designs support dose optimization, patient stratification, trial enrichment, and the evaluation of combination therapies. As such, Breakthrough T1D is generally supportive of adaptive designs which can help accelerate the development and availability of new therapies that meet the significant unmet needs faced by those living with T1D. The ICH draft guidance would benefit by including some of the unique considerations seen for trials in the cell therapy field, as trials of these products are often single-arm and unblinded. While adaptive trials use interim analyses of results to modify a trial according to predefined rules, in the cell therapy field mid-trial changes may be more likely that include refining the sample size, eliminating treatments or doses, identifying patients who are most likely to benefit in order to focus recruitment, or terminating a trial because of clear success or failure. It would be valuable for EMA to expand on the ICH guidance to provide additional guidance on this topic to support developers in designing more practical and scientifically robust trials. Such guidance could help ensure that adaptive designs remain applicable and valuable even in these unique and resource-constrained settings.	
Breakthrough T1D	0	0		Finally, developers would benefit from illustrative examples within the final guidance or accompanying materials that demonstrate scenarios where adaptive trial designs can be effectively applied. Such examples, particularly if real world examples are highlighted, would enhance clarity and facilitate implementation.	
EFPIA	0	0	0	The guidance provides limited novelty as compared to previous guidance on adaptive designs. While one could interpret the current draft as an “open regulatory position” towards implementation of adaptive designs (if key principles are followed), specifics provided in subsequent sections significantly raise concern on regulatory acceptability. In particular, the draft guidance is repetitive in demanding justification for the adaptive design and is thereby providing plenty opportunities for unjustified mechanisms to push back in situations where still “key principles” are followed (e.g. lines 891/892: “evaluate acceptability of any additional uncertainty attributable to proposed adaptive elements”). The ICH expert working group should have a general understanding that there is no complete certainty on assumptions during the planning stage – otherwise, a confirmatory trial might not be required, as everything is already known. Adaptive designs are frequently implemented to address uncertainties and certainly not to cut corners. One may rather need to reconsider, whether designs without adaptive elements should require justification. Clinical trials without interim analyses run the risk that one would just learn at a final analyses that one should have terminated the trial early for futility and could have thereby treated patients with a better treatment and decreased resource spending (patient’s time and efforts, as well as research budgets). The guidance currently will build new hurdles by introducing additional unnecessary and time-consuming “documentation requests” (e.g. simulation and documentation), for which the added utility is not always obvious (e.g. situations such as GSD or Futility Stopping). While it is understood that a differentiation of “well understood” vs. “less well understood” designs is not desired anymore in 2025, there current draft guidance unfortunately appears to recognize all adaptive designs as “less well understood”.	The guidance requires significant reconsideration to limit the amount of excess work in situations, where adaptive designs are well understood. It needs to become more specific on situations, where adaptive designs are considered conditionally acceptable and where additional justification is required. It needs to identify what conditions need to be discussed with regulators to accept the adaptive design. Sponsors should not be constantly challenged to justify adaptive designs. It is in everyone's interest to find treatment effectiveness and safety as early as possible. Patients are waiting. In many cases, a non-adaptive design should be justified and challenged as well. It would be useful to highlight in parts of the document that GSD is a special type of adaptive design that is well-understood and commonly used in practice.
EFPIA	0	0	0	The focus of the guideline appears to be inferentially seamless designs. This makes sense as these have the greater opportunities for issues in terms of bias, data integrity etc. However, operationally seamless designs are possible but are not acknowledged at all in the draft.	Add a statement around focus on inferentially adaptive designs in the "Introduction and Scope".
EFPIA	0	0	0	Overall, this guideline is comprehensive and clearly written. It provides clear and practical recommendations that are useful for sponsor to specify and document an adaptive design, and engage collaborative discussions with regulators. Comments provided in this document are for consideration.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	0	0	0	Several places references are made to benefit-risk evaluations and how they might cause deviations from strict pre-specified statistical rules. Such events which would typically via a iDMC evaluation are difficult to take into account in the prospective planning of the trial design and can't as such be assesed at the planning stage via simulations. Is it acceptable to acknowledge up front that these events can occur but not formally account for them in the statistical plans, but rather adjust if they occur?	
EFPIA	0	0	0	The explanatory note on the title page acknowledges the "high potential for adaptive designs to accelerate the process of drug development and to allocate resources more efficiently without lowering scientific and regulatory standards". It seems that while the intentions of this guidance are well-defined, this is not obvious from the remainder of the text, which often challenges applicability of adaptive designs. References to the high potential of adaptive designs and their positive aspects with regard to time, effectiveness and feasibility are largely missing in the main text.	Add examples of the high potential of adaptive designs and their positive aspects with regard to time, effectiveness and feasibility in the main text.
EFPIA	0	0	0	It is surprising that missing data is not mentioned as a factor to consider when comparing adaptive designs. This could be a source of uncertainty and one that if affecting variables of adaptation can in turn affect the operating characteristics by which a particular design (adaptive or not) is chosen. At the very least, this topic should be mentioned as a factor to be considered as part of choosing and justifying an adaptive design.	Suggest adding a line mentioning the consideration of missing data affecting outcome data that would be use to trigger adaptations in the section 5.1 (and/or in the general principles for adaptive designs) and emphasise the importance of implementing the estimand focused imputation methods not just at the end of the trial but at interim stages if needed and as part of the design proposed.
EFPIA	0	0	0	No reference provided	Please add references to highlight which statistical methods are of interest (non-exhaustive list).

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	0	0	0	The draft ICH E20 Harmonised Guideline entitled "Adaptive Designs for Clinical Trials" addresses a substantial breadth of considerations in the evaluation of a confirmatory clinical trial. We applaud the tremendous effort towards clarifying the considerations surrounding the design and implementation of adaptive clinical trials in the confirmatory setting. While the draft Guideline addresses these considerations in the context of the design, implementation, and analysis of adaptive clinical trials, many of these considerations apply similarly to traditional frequentist, non-adaptive confirmatory trials as well. This document has value in supporting the appropriate, productive, and careful application of adaptive clinical trial design. However, the ability to maintain trial integrity with concomitant improvements in successfully identifying safe and effective treatments will be limited by general factors listed below: (1) Insufficient focus on the consistent and equal application of objective criteria to the evaluation of both adaptive and non-adaptive trials. This in turn is associated with a lack of objective judgement of non-adaptive approaches against equivalent benchmarks. For instance, there is little or no acknowledgement of the limitations of and risks associated with the use of traditional, non-adaptive approaches to trial design (e.g., the risks of continuing a trial longer than necessary without interim analyses and prespecified futility rules, or in conducting a large simple trial to estimate an average treatment effect in a population in which heterogeneity of treatment effect is likely, resulting in a highly-precise estimate of a treatment effect that applies to no one). (2) A general representation throughout the document that adaptive designs are at risk of providing less information than non-adaptive clinical trial designs. While this may be true if the adaptive design is poorly designed—any approach can be done poorly—the document should take the opportunity both to clarify the characteristics of high-quality adaptive design and to compare adaptive and non-adaptive approaches under the assumption that both are well designed and executed. The motivation of adaptive designs is to strive for the right amount of data to meet pre-defined user requirements or desired operating characteristics. This applies to stopping just at the right time (including increasing the sample size to provide additional information in light of the accruing data), response-adaptive randomization to maximize the information on the arm found to be of greatest interest, and other adaptations.	(1) The guideline should explicitly emphasize the equal application of objective criteria to the both adaptive and non-adaptive trials, including a balanced summary of the risks of non-adaptive approaches from patient, sponsor, and regulatory perspectives. (2) The implication that adaptive trial designs provide less information useful in regulatory decision making should be eliminated and replaced with statements that the infomation obtained from each approach is highly dependent on the specifics and quality of the design, execution, and analysis for each approach.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	0	0	0	[Continued from Above] (3) There is a missed opportunity to emphasize the importance of quantification of threats to trial validity (e.g., bias in estimation) as well as a consistent emphasis on certain threats (e.g., bias, type I error, risks to trial integrity) without an equivalent consideration of the other risks (e.g., variance in estimation, failing to consider valid external evidence, reduced power or ability to identify the most effective dose or adverse safety signals). This is accompanied by another missed opportunity, namely to discuss the advantages of some estimation strategies (e.g., hierarchical models) and trial designs (e.g., basket trials) that demonstrate improved performance by important metrics. One specific omission is a discussion of the biases associated with common, traditional approaches, e.g., the upward bias seen in the largest among many raw estimates. (4) An implied equivalence between the use of Bayesian strategies and the borrowing of external information. In contrast, there are non-Bayesian approaches to borrowing external information, Bayesian methods are most commonly used without substantial external information, and the topic of borrowing deserves a stand-alone guideline. (5) Inconsistent emphasis on the importance of adhering to prespecified rules for adaptations in maintaining the defined operating characteristics of an adaptive trial, e.g., use of the term “anticipated” rule (e.g., on lines 127, 145-154, 318, 329, 388, 394, 414-418, and following) or suggesting a routine role for an IDMC in determining whether or how a prespecified rule is applied as part of the design (e.g., see lines 145-148). This results in a lack of clarity in whether the design is thoroughly prespecified, or represents simply a range of possible options regarding trial implementation. There are missed opportunities associated with the strategy of maintaining flexibility in the application of adaptive rules: To maintain appropriate operating characteristics such as type I error control may require the use of such conservative analysis strategies that the desired efficiency of the adaptive design is lost. This tradeoff should be discussed.	(3) The guideline should emphasize the importance of quantification of threats to trial validity (e.g., bias in estimation), and balance the discussion of important threats (e.g., bias, type I error, risks to trial integrity, variance in estimation, failing to consider valid external evidence, reduced power or ability to identify the most effective dose or adverse safety signals). The guideline should mention the advantages of some estimation strategies (e.g., hierarchical models) and trial designs (e.g., basket trials) that demonstrate improved performance by important metrics. Finally, the guideline should acknowledge that biases exist in some common, traditional approaches, e.g., the upward bias seen in the largest among many raw estimates. (4) A distinction needs to be made between the use of Bayesian strategies and the borrowing of external information. The guideline should acknowledge that there are non-Bayesian approaches to borrowing external information and that Bayesian methods are most commonly used without substantial external information. The topic of borrowing is complex and likely deserves a stand-alone guideline. (5) The guideline should strive to achieve a consistent emphasis on the importance of adhering to prespecified rules for adaptations to maintain defined operating characteristics of an adaptive trial and acknowledge the risk of allowing aroutine role for an IDMC in determining whether or how a prespecified rule is applied as part of the design. Similarly, the guideline should mention the disadvantages associated with the strategy of maintaining flexibility in the application of adaptive rules.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	0	0	0	[Continued from Above] (6) There is a missed opportunity to acknowledge the ethical imperatives (i) to optimize statistical efficiency and expose as few participants as possible to experimental treatments in determining safety and efficacy; and (ii) to minimize the risk of failing to correctly identify treatments that are safe and effective, i.e., type II errors. The focus on type I error, while critically important, fails to address regulatory agencies' broader remit to optimize population health, which is not limited to simply preventing ineffective treatments from reaching the market. (7) There is a pervasive implication that adaptive designs result in smaller enrolled populations and are less informative—a shortcut. There are many examples where, in practice, adaptive designs and advanced modeling provide more information and better precision regarding parameters that are critical to regulatory decision making. Three simple examples: (i) if forced to do a fixed sample size in a 1:1 confirmatory trial, a smaller sample size is generally selected, while the use of interim analyses to test for superiority and futility allow trial to be larger than the typical fixed trial when needed, but only when needed. Adaptive-sample-size phase 3 trials are almost always more powerful than fixed-sample-size phase 3 trials; (ii) a seamless 2/3 trial that selects dose(s) to continue to a confirmatory stage will be larger, as the data used for doses selection is included in the confirmatory analysis, this tends to lead to larger phase 2 trials with better dose characterization; and (iii) if a basket or enrichment trial is conducted this allows the ability to better characterize heterogeneity of the treatment effect or safety profile, resulting in a better selected population. In practice, the alternative non-adaptive strategy is virtually never a wide-ranging and adequately sized phase 2 trial of the broader population, it is a selection of a single population with the associated risks to both the population and the sponsor, followed by the conduct of a non-adaptive large phase 3 trial. Instead of communicating the current perspective that has the potential, perhaps paradoxically, to decrease the safety and informativeness of trials conducted to inform product development and regulatory decision making, the draft guideline has an opportunity to discuss the potential value of appropriately used adaptive design elements to substantially improve clinical drug development and confirmatory clinical trials.	(6) The guideline should acknowledge the ethical imperatives (i) to optimize statistical efficiency and expose as few participants as possible to experimental treatments in determining safety and efficacy; and (ii) to minimize the risk of failing to correctly identify treatments that are safe and effective, i.e., type II errors. (7) The implied assumption that adaptive trial designs uniformly result in smaller enrolled populations or provide less information should be replaced by a statement that the size of the enrolled population and the precision of important results depends on the details of the design and associated decision rules. Examples as provided in the comment should be used to clarify this point. The guideline should acknowledge the potential value of appropriately used, high-quality adaptive design elements to substantially improve clinical drug development and confirmatory clinical trials.
EFPIA	0	0	0	In Section 5.3, the Bayesian approach is largely equated to the use of informative prior distributions for borrowing of information and, perhaps, even to a static approach to borrowing of information. However, in the vast majority of applications of Bayesian methods in adaptive clinical trial design, relatively non-informative prior information is used, prior information that is rapidly overwhelmed by accumulating data. The benefits of the Bayesian approach in this setting include the coherent inferential framework and interpretability of adaptive rules based on posterior probability distributions or, unique to Bayesian inference, predictive probabilities. Many modern implementations of Bayesian borrowing of information utilize a dynamic approach in which the degree of borrowing depends on the observed consistency in treatment effect between the borrowed and newly acquired data. This approach can largely mitigate risks associated with static borrowing approaches and should be explicitly mentioned as a strategy that is worthy of consideration. Such an approach often requires careful selection of the prior variance on the distribution of treatment effects, a design choice that should generally be supported by simulation studies. The discussion of type I error control in the setting of the use of external information will likely be misleading to many readers. When appropriate information is borrowed, this is inferentially equivalent to the "seeding" of a traditional standalone trial with an initial set of participants. If those data are consistent with efficacy of the experimental treatment, then the probability of a positive conclusion, if the new data arise from a population in which there is no treatment benefit, should not be controlled at the usual type I error rate. An increase in the "nominal" type I error risk reflects, in essence, the information and value of the external information. Maintaining traditional type I error control would require making the criteria for demonstrating efficacy sufficiently stringent to neutralize the positive effect of the promising borrowed information, an approach that would defeat the intended purpose of the design.	As mentioned above, the discussion of the Bayesian approach should be separated from the general discussion of the borrowing of information. The general use of Bayesian inference with minminally informative prior distributions should be acknowledge, along with the potential advantages of this approach. The value and advantages of dynamic borrowing strategies, whether Bayesian or frequentist in application, should be mentioned. The discussion of type I error in the setting of the use of external information needs to be reformulated to eliminate the implication that the type I error rates should be controlled at traditional levels with the incorporation of external information. The concept of type I error in the setting of supportive external information needs careful definition and discussion as, in essence, the type I error rate of interest is the rate for the entire procedure including the generation of the external information, rather than the type I error rate conditional on the external data that were ultimately observed.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	0	0	0	[Continued from Above] As currently written, the draft Guideline could be interpreted as suggesting the need for nominal type I error control, e.g., 0.025 one-tailed, in the current trial with the inclusion of borrowing. We agree that the probability of a positive trial result if all new data arise from a population experiencing no treatment benefit should be quantified, e.g., through simulation, and understood; however, the acceptable rate of a positive trial result in this context should depend on the details of the clinical and regulatory setting. The discussion of Bayesian approaches to adaptive design needs to be separate from the discussion of borrowing of information; while Bayesian approaches are well suited for sophisticated approaches to borrowing (e.g., dynamic borrowing), these are separate concepts and frequentist methods of borrowing should also be acknowledged. The draft Guideline focuses on trial design in the confirmatory setting and, in that context, the inclusion of Section 5.5 on exploratory trials seems out of place and confusing. The role of adaptive design in the exploratory or learn-phase setting is well established and the balance between the needs for flexibility, efficiency, control of type I error, and integration of efficacy and safety considerations (e.g., in dose selection) are all quantitatively and qualitatively different than in a confirmatory setting. The draft Guideline risks a false equivalence, implying that exploratory trials should adhere to the requirements for confirmatory trials. While many of the general points about steps to ensure trial validity apply across these two settings, those considerations are generally not specific to adaptive designs and are covered elsewhere in ICH Guidelines and other regulatory guidance documents. We strongly recommend that the discussions of exploratory trials be removed from this Guideline, with an explicit statement that the associated considerations are discussed elsewhere.	See above regarding the definition and control of type I error risk in the setting of supportive external information. The section on exploratory trials should be removed from this guideline to avoid implying a false equivalence that exploratory trials should adhere to the same requirements as confirmatory trials.
EFPIA	0	0	1	The material presented in the draft guideline is challenging for non-statisticians, as a statistical background is required to fully comprehend much of the content. The development of appropriate training materials would facilitate better understanding among readers, which in turn would facilitate more effective discussion and collaboration within sponsor organisations.	Please see suggestion to develop additional training materials for non-statisticians to facilitate internal discussions.
EFPIA	0	0	2	The guidance stresses the need for “justification for an adaptive design”. Any clinical trial, be it adaptive or non-adaptive, complex or simple should be thoroughly designed to assert efficiency and robustness in decision making. Interim analyses provide best opportunities for increasing efficiency and should hence be conventionally evaluated. As such, the need for justification should not be limited to “adaptive designs”, but “justification” should equally be required for non-adaptive designs, which for example do not implement interim looks assessing futility.	
EFPIA	0	0	3	The guideline recommends, very reasonably, that adaptive trial designs should not be overly complex and adaptation should occur at a single, clearly defined point. This is fine as long as it is made clear what type of adaptations are referred to here. In particular, a good group sequential design should have several interim analyses to allow early stopping, either for efficacy or futility. Table 4.4 of Jennison & Turnbull (1999) shows the efficiency gains that can be achieved by group sequential designs: these increase with the number of interim analyses. The benefits of an additional analysis gradually decrease and these authors recommend designs with 4 or 5 interim analyses.	Reference: Jennison, C. and Turnbull, B. W. (1999) Group Sequential Methods with Applications to Clinical Trials. Chapman & Hall/CRC.
EFPIA	0	0	3.4	It is good that the word "limited" appears in the requirement for "limited or no bias in the primary estimate of treatment effect". Emerson and Fleming (1990) showed it is possible to obtain a minimum variance unbiased estimate of the treatment effect after a group sequential trial, but it is well known that such estimates have a high variance. Whitehead (Biometrika, 1986) proposed a simple method to obtain an adjusted maximum likelihood estimate (adjusted MLE). Typically, this estimate has some bias, but the amount of bias is negligible for practical purposes and mean square error is smaller than that of an unbiased estimate.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	0	0	3.4	In the case of a 2-stage group sequential design, the minimum variance unbiased estimate has peculiar properties. If the trial stops at analysis 1 with rejection of the null hypothesis, one might think it appropriate to adjust the maximum likelihood estimate downwards to allow for "stopping on a random high". However, the unbiased estimate makes no adjustment in this case. If the trial stops at analysis 2 with a very clear rejection of the null hypothesis, the unbiased estimate makes a big downwards adjustment which can be much larger than the maximum bias of the MLE at any value of the true treatment effect. In view of these facts, one could modify the requirement by removing the words “no bias” and ask for "limited bias in the primary estimate of treatment effect".	References: Emerson, S.S. and Fleming, T.R. (1990). Parameter estimation following group sequential hypothesis testing. Biometrika, 77, 875–892. Whitehead, J. (1986). On the bias of maximum likelihood estimation following a sequential test. Biometrika, 73, 573–581. https://people.bath.ac.uk/mascj/talks_2025/cj_slides_EFSPI.pdf
EFPIA	0	0	3.4	The guidance repeatedly stresses concerns on biased estimates. It is understood that point estimates are of relevance for benefit-risk decision making. Still, the focus on bias diverts attention from making decisions efficiently, what is particularly relevant to the patients who are taking part in the clinical trials, as well as those outside of the trial. While there is bias introduced through adaptive designs, also the magnitude and direction will need to be considered. Minor concerns on bias should not outweigh the efficiencies generated with adaptive designs.	
EFPIA	0	0	4	The early focus on the key principles helps emphasise the minimum requirements on any adaptive design, when it comes to generating confirmatory evidence is compelling. Also, the second section is clear in summarizing most frequently considered adaptations. Considerations on other types of adaptations should be included here, emphasizing potential important considerations, if those are less well understood. In particular descriptions regarding “choice of endpoint or testing strategy”, which could also include adaptive changes from non-inferiority to superiority would be of relevance, change in imaging modality, change in spending function, etc. . Considerations for platform, basket and umbrella trials are missing. Those designs have been emerging in the past years as promising design candidates for generating evidence in situations of unmet need, such as in Covid-19 and for rare indications.	Add to the section on "Types of adaptations" (Section 4), subsections on "Change of testing strategy", "Adding trial arms" and "Platform, Basket and Umbrella trials".
EFPIA	0	0	4	There are designs, which are well understood and designs, which are less well understood, even if those or related terms did raise controversies in the past. It would be beneficial if well understood adaptations (futility, GSD, sample size re-estimation), would be considered as “standard” and would be more explicitly stated as being generally acceptable. It is acknowledged that justification of the chosen adaptation rules will still be required.	
EFPIA	0	0	4.2	In the discussion regarding blinded vs. unblinded sample size reassessment the risk of false adaptations due to blinded sample size reassessment should also be discussed, which could be avoided by having the IDMC performing an unblinded look.	
EFPIA	0	0	4.2	Guidance on adaptive sample size decrease should be provided.	
EFPIA	0	0	4.2	The guidance reflects type 1 error a lot which is understandable, however it ignores the fact that decision errors like type I and II errors have multiple meanings in certain adaptive designs such as blinded sample size re-estimation for non-inferiority and equivalence trials.	Suggest the guidance to add a statement about such nuances and the need for researchers to clarify and justify the decision errors they are controlling in such setting
EFPIA	0	0	4.4	The guidance ignores adaptive designs for umbrella or platform trials settings. Addition of new arms to an ongoing trial is an adaptive method, if driven by interim accumulated data. Unsure if this was intentional or if a new ICH topic adoption is being planned for master protocol trials specifically.	Considerations for adaptive methods (such as RAR or drop the loser designs) for Multi-Arm Multi-Stage (MAMS) and Platform Trials should be included in section 4.4 (Treatment Selection)
EFPIA	0	0	5	CT with adaptive design may provide limited safety data as acknowledged in the document.	In addition to the proposals already included, the document could suggest the Sponsors perform aggregate safety data review at the program level (if available) to inform the B:R assessment at a given single CT with adaptive design

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	0	0	5	A statement on considerations related to the use of RWD, including relevant references that readers should consider, would be useful. We acknowledge that material changes to the document are not possible at this stage and, if appropriate, this topic may eventually be addressed in ICH E23.	Propose that Section 5 would benefit from covering the use of external data in an adaptive design. One example is borrowing external data to augment the control arm in a randomized trial. The adapatation could be increasing the sample size if the chance of borrowing external data is low based on the pre-specified criteria.
EFPIA	0	0	5	Add to "Special topics and considerations" information on non-prospectively planned adaptations.	Add to "Special topics and considerations" information on non-prospectively planned adaptations; Discuss in section 5 adaptations to sources external to the trial and when they can or cannot anymore be implemented.
EFPIA	0	0	5.2	The very detailed requirements for the conduct and reporting of simulation studies stand out from the rest of the document and may not be appropriate in all cases. We suggest allowing a more nuanced approach depending on the situation and the purpose of the simulation study. Rigor and detail will be required for controlling false benefit/risk decisions in confirmatory settings. Other situations however may require less, e.g. if the main decision probabilities are protected by theory, and the simulation merely serves ancillary purposes (tertiary objectives, operational forecasting etc.).	
EFPIA	0	0	5.3	Bayesian methods can be efficiently implemented in confirmatory trials for rare diseases due to limited sample size. The guidance document speaks much about external borrowing of information as a specific case to a Bayesian adaptive trial ut misses out on how these methods can be specifically implemented for rare disease setting.	Suggest including a discussion on implentation of Bayesian methods in rare disease confirmatory trials as a special case of the use of adaptive design elements in the context of clinical trials that use Bayesian methods.
EFPIA	0	0	5.3	Section 5.3 (Adaptive Designs Using Bayesian Methods) describes the considerations and potential caveats for borrowing information from external data. Since there are also non-Baysesian methods (Frequenists' methods) proposed for borrowing information from external data, it will be useful to clarify whether the considerations and caveats discussed in this section are limited to Bayesian approaches only, or they are applicable for all methods that are borrowing information from external data (regardless of whether they are Bayesian or Frequentists' methods).	If the considerations are also applicable to non-Bayesian methods, consider adding a sentence at the end of the section for clarification, for example, "The considerations discussed in this paragraphs are applicable to all types of methods that borrow information from external data, including the Bayesian approaches discussed in this section, as well as the Frequentists' approaches developed recently." Reference: Li R, Lin R, Huang J, Tian L, Zhu J. A frequentist approach to dynamic borrowing. Biom J. 2023 Oct;65(7):e2100406. doi: 10.1002/bimj.202100406.
EFPIA	0	0	5.3	The section on Bayesian Approaches is not introducing a discussion on the relative weight of the prior distribution in interim analyses. There will be less data available at interim analyses, such that the prior would be expected to have a greater weight on trial read-outs in interim analyses (vs. final analyses). It may be important to name this issue, such that it could be properly addressed in designing trials.	
EFPIA	0	0	5.3	The footnote is unnecessarily discouraging and does not provide any justification. The "Key Principles" outlined in Section 3 should be the same for frequentists vs. Bayesian approaches. Thus, it is not directly clear, why such a footnote on Bayesian approaches is necessary. If kept, it should be made clearer which key principles are supposedly impacted by Bayesian approaches.	Remove footnote and fully harmonise section on "Adaptive Designs Using Bayesian Methods".
EFPIA	0	0	5.3	In discussing Bayesian methods, the guideline avoids using the term "type I error rate". It does, however, say that it is important to limit the chances of erroneous conclusions. A more explicit statement of what this requirement means would be helpful.	In discussing Bayesian methods, the guideline avoids using the term "type I error rate". It does, however, say that it is important to limit the chances of erroneous conclusions. A more explicit statement of what this requirement means would be helpful.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	0	0	5.3	Once a Bayesian analysis has been defined, it has a Type I error rate (or a familywise error rate if multiple hypotheses are tested). This is simply a conditional probability given that the treatment effect is zero (or the maximum probability of a false positive over a set of scenarios). In a Bayesian analysis, one may be interested in other probabilities, a marginal probability obtained by integrating over a prior distribution, or a conditional probability given a posterior distribution, and a Bayes decision rule may be defined in terms of these. However, when assessing the operating characteristics of a Bayesian design, and applying the same criteria as for other types of design, it is natural to consider the type I error rate. Indeed, "calibrated" Bayes designs do this when setting criteria for decision making.	
EFPIA	0	0	5.3	There is a need to increase clarity on Bayesian designs vs. Borrowing. Non-Bayesian approaches to borrowing do exist and are increasingly discussed in context of external controls using causal inference methods. If concerns apply to any type of borrowing, a separate section on Borrowing should be considered to avoid confusion with Bayesian design proposals. Increased clarity is also required on the requirement for "patient-level data" (line 767). "Patient-level data" will frequently not be available (as opposed to aggregate data), such that the guidance appears to effectively exclude most applications of borrowing.	
EFPIA	0	0	5.6	Pausing enrollment in some multi-stage adaptive designs could be an operational challenge, especially when enrollment is rapid but the endpoint for decision-making in the earlier stage requires a long follow-up time.	Consider acknowledging the operational challenges to resume enrollment for the later stage of the trial after the pause, and indicate the preference for designs that don't require pausing enrollment.
EFPIA	0	0	6	It is not entirely clear from the guidance, why a separate section on Documentation is required – and if required, information on the appropriate place to document the design should be provided to prospectively harmonize the approach for document submission.	Prospectively harmonize the approach for document submission for ICH regarding adaptive designs.
EFSPI	0	0	2	The guidance stresses the need for "justification for an adaptive design". Any clinical trial, be it adaptive or non-adaptive, complex or simple should be thoroughly designed to assert efficiency and robustness in decision making. Interim analyses provide best opportunities for increasing efficiency and should hence be conventionally evaluated. As such, the need for justification should not be limited to "adaptive designs", but "justification" should equally be required for non-adaptive designs, which for example do not implement interim looks assessing futility.	
EFSPI	0	0	3.4	The guidance repeatedly stresses concerns on biased estimates. It is understood that point estimates are of relevance for benefit-risk decision making. Still, the focus on bias diverts attention from making decisions efficiently, what is particularly relevant to the patients who are taking part in the clinical trials, as well as those outside of the trial. While there is bias introduced through adaptive designs, also the magnitude and direction will need to be considered. Minor concerns on bias should not outweigh the efficiencies generated with adaptive designs.	
EFSPI	0	0	4	There are designs, which are well understood and designs, which are less well understood, even if those or related terms did raise controversies in the past. It would be beneficial if well understood adaptations (futility, GSD, sample size re-estimation), would be considered as "standard" and would be more explicitly stated as being generally acceptable. It is acknowledged that justification of the chosen adaptation rules will still be required.	
EFSPI	0	0	5.2	The very detailed requirements for the conduct and reporting of simulation studies stand out from the rest of the document and may not be appropriate in all cases. We suggest allowing a more nuanced approach depending on the situation and the purpose of the simulation study. Rigor and detail will be required for controlling false benefit/risk decisions in confirmatory settings. Other situations however may require less, e.g. if the main decision probabilities are protected by theory, and the simulation merely serves ancillary purposes (tertiary objectives, operational forecasting etc.).	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFSPI	0	0	5.3	There is a need to increase clarity on Bayesian designs vs. Borrowing. Non-Bayesian approaches to borrowing do exist and are increasingly discussed in context of external controls using causal inference methods. If concerns apply to any type of borrowing, a separate section on Borrowing should be considered to avoid confusion with Bayesian design proposals. Increased clarity is also required on the requirement for "patient-level data" (line 767). "Patient-level data" will frequently not be available (as opposed to aggregate data), such that the guidance appears to effectively exclude most applications of borrowing.	
EFSPI/PSI Regulatory ESIG	0	0		According to the overall impression, the effort to justify adaptive designs seems much higher that the effort for justifying a non-adaptive design. (As an example, documentation of simulations around designs with a futility analysis or standard GSD may be perceived as increasing the bar.)	Consider presenting risks and benefits associated with non-adaptive and adaptive designs more neutrally and, perhaps, point out that in same situations not implementing an adaptive design may be detrimental.
EFSPI/PSI Regulatory ESIG	0	0	6.6	It would be helpful to draw attention to the recruitment of patients between the date the data cut is taken and the date when decisions are made based on interim analysis results, to ensure a plan is made early in the study. For example, should recruitment be paused during this period, how will the data for such patients be handled in the CSR analyses?	Add text: Consideration should be given to the recruitment of new patients between the date the interim analysis data cut is taken and the date when decisions are made based on interim analysis results. Consider if it is in the patient's best interest to be recruited to an arm that might be dropped due to futility once the interim analysis results become available. Such patients will not contribute to the interim analysis and might not contribute to the primary/secondary endpoint analyses if the arm which is dropped is excluded from primary/secondary analyses.
EFSPI/PSI Regulatory ESIG	0	0	6.6	It would be helpful to mention special considerations for the Statistical Analysis Plan, to reduce the risk of making snap decisions at the time of interim analysis which could have serious impact on results.	Add text: The Statistical Analysis Plan should be used to carefully describe additional data handling and analysis consideration specific to the interim analysis. For example, it should be clear if all data in the database at the time of data cutoff should be included in analyses (cumulative), or if data should be restricted to patients who have completed/discontinued prior to the specified timepoint (complete). Different strategies might be necessary for efficacy and safety analyses. Key dates such as treatment end date or adverse event date will often be missing at the time of interim analysis and appropriate imputation rules should be described.
EFSPI/PSI Regulatory ESIG	0	0	6.6	Changing the sample size during the study can have downstream impact on processes not related to the statistical analyses. It would be useful to draw attention to this so that plans to mitigate any potential risks related to this are put in place at the appropriate time.	Add text: Ensure appropriate attention is given to all study plans impacted by a change in sample size, so that they are updated in a timely manner following changes to sample size. For example, a change in randomization probabilities will impact randomization schedules and a plan is necessary to ensure the blind is maintained (if necessary); a study plan describing Quality Tolerance Limits (ICH E6 (R3)) will be impacted by a change in planned sample size if QTL calculations and reporting include the planned total sample size for the study.
EORTC	0	0	0	There appears to be a lack of explicit examples or references to examples, that would help illustrate some of the key principles. Although the scope notes that the guideline does not discuss specific statistical methods, including a few basic references, without elaboration, would be very helpful for readers.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
c4c-S	0	0		Overall	The draft discusses data latency but does not provide criteria for determining acceptable data lag in adaptive decision making.
c4c-S	0	0		Overall	Although timely data cleaning is emphasized, the document does not specify minimum quality thresholds required before interim analyses.
c4c-S	0	0		The guideline appropriately connects adaptation triggers to clinical plausibility but would benefit from clearer criteria from defining such triggers.	1. Explicit minimum documentation standards for simulation studies and IDMC communication. 2.More precise expectations for Bayesian calibration and transparency. 3.Formal reproducibility and software validation requirements for simulation and adaptive algorithms. 4. Greater clarity on sponsor governance boundaries to preserve trial integrity.
c4c-S	0	0	3.1-9	Regarding the rest of proposed draft to EMA	In agreement with Prof. Lithoxopoulou Maria.
Ferring Pharmaceuticals	0	0		Could a definition of "Interim Analysis" be included, e.g. like in the FDA guidance on adaptive designs? It could be helpful to try to standardize terminology. If there is an interest to further standardize terminology, suggestions from https://arxiv.org/abs/2410.01478v2 could also be considered.	
IDSWG	0	0	4.4	This section shouild at least mention the selection of treatments from adaptive platform trials	Add text mentioning adaptive platform trials with concurrent and non-concurrent controls and their utilization for selection of treatments
IDSWG	0	0	5.3	The Bayesian Adaptive Design is incomplete without the mentioing of Platform Trial Designs where Bayesian methods are most suitable when different investigational treatments enter the adaptive trial at different points in time. This also sets up a scenario where the borrowing of information, especially for the control group could occur within the scope of the ongoing clinical trial either concurrently or non-concurrently	A paragraph should be dedicated to the use of Bayesian methods for decision rules and the borrowing of control group information in Master Protocols (Basket, Umbrella, and Platform Trials).
IDSWG	0	0	5.3	Section 5.3 (Adaptive Designs Using Bayesian Methods) describes the considerations and potential caveats for borrowing information from external data. Since there are also non-Baysesian methods (Frequentists methods) proposed for borrowing, information from external data, it will be useful to clarify whether the considerations and caveats discussed in this section are limited to Bayesian approaches only, or they are applicable for all methods that are borrowing information from external data (regardless of whether they are Bayesian or Frequentists methods).	If the considerations are also applicable to non-Bayesian methods, consider adding a sentence at the end of the section for clarification: “The considerations discussed in this paragraphs are applicable to all types of methods that borrow information from external data, including the Bayesian approaches discussed in this section, as well as the Frequentists approaches developed recently.” Reference: Li R, Lin R, Huang J, Tian L, Zhu J. A frequentist approach to dynamic borrowing. Biom J. 2023 Oct;65(7):e2100406. doi: 10.1002/bimj.202100406
IDSWG	0	0	5.6	Pausing enrollment in some multi-stage adaptive designs could be an operational challenge, especially when enrollment is rapid but the endpoint for decision-making in the earlier stage requires a long follow-up time.	Acknowledge the operational challenges to resume enrollment for the later stage of the trial after the pause, and indicate the preference for designs that don’t require pausing enrollment.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
IDSWG	0	0	5.3	The Bayesian Adaptive Design is incomplete without the mentioing of Platform Trial Designs where Bayesian methods are most suitable when different investigational treatments enter the adaptive trial at different points in time. This also sets up a scenario where the borrowing of information, especially for the control group could occur within the scope of the ongoing clinical trial either concurrently or non-concurrently	A paragraph should be dedicated to the use of Bayesian methods for decision rules and the borrowing of control group information in Master Protocols (Basket, Umbrella, and Platform Trials).
IDSWG	0	0	5.3	Section 5.3 (Adaptive Designs Using Bayesian Methods) describes the considerations and potential caveats for borrowing information from external data. Since there are also non-Baysesian methods (Frequentists methods) proposed for borrowing, information from external data, it will be useful to clarify whether the considerations and caveats discussed in this section are limited to Bayesian approaches only, or they are applicable for all methods that are borrowing information from external data (regardless of whether they are Bayesian or Frequentists methods).	If the considerations are also applicable to non-Bayesian methods, consider adding a sentence at the end of the section for clarification: "The considerations discussed in this paragraphs are applicable to all types of methods that borrow information from external data, including the Bayesian approaches discussed in this section, as well as the Frequentists approaches developed recently." Reference: Li R, Lin R, Huang J, Tian L, Zhu J. A frequentist approach to dynamic borrowing. Biom J. 2023 Oct;65(7):e2100406. doi: 10.1002/bimj.202100406
IDSWG	0	0	5.6	Pausing enrollment in some multi-stage adaptive designs could be an operational challenge, especially when enrollment is rapid but the endpoint for decision-making in the earlier stage requires a long follow-up time.	Acknowledge the operational challenges to resume enrollment for the later stage of the trial after the pause, and indicate the preference for designs that don't require pausing enrollment.
c4c-S	0	0		Adaptive designs are particularly relevant for paediatric development programmes where small sample sizes, ethical considerations, and heterogeneity between age groups make flexibility crucial. The guideline could more explicitly address paediatric contexts, including age-stratified subcohorts, extrapolation strategies, and data borrowing from adult studies.	
c4c-S	0	0	3	While Section 3 provides a comprehensive framework for maintaining the scientific validity and interpretability of adaptive confirmatory trials, several provisions remain overly qualitative.	Key areas needing enhancement include:Quantitative standards for simulation validation, Type I error control, and bias evaluation. Explicit criteria for permissible deviations from adaptation rules. Operational integrity metrics for information control and IDMC governance. Greater inclusivity of Bayesian and complex adaptive designs with simulation-supported validation.
c4c-S	0	0		Section 5 effectively consolidates key operational and methodological considerations but would benefit from	1. Explicit minimum documentation standards for simulation studies and IDMC communication. 2.More precise expectations for Bayesian calibration and transparency. 3.Formal reproducibility and software validation requirements for simulation and adaptive algorithms. 4.Greater clarity on sponsor governance boundaries to preserve trial integrity
c4c-S	0	0		Overall	Include a formal Adaptive Design Documentation Checklist for both pre-trial and post-trial submissions.
c4c-S	0	0		Overall	Add clarity on hybrid Bayesian–frequentist contexts and handling of prior-data conflicts.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
c4c-S	0	0	3	Overall in agreement with the current text. It is indeed rather qualitative, but making it more quantitative can be impractical when it is to be used in daily practice.	No change
EFPIA	0	0	5.2	The section on simulation studies starts with a "often play an important role". It would be very helpful to clarify for which situations simulation studies are required, as the section might be easily misunderstood that every adaptive design needs to be accompanied by a simulation report covering "all plausible scenarios". If this section is followed literally, it would add frequently non-required burden to the design and review process, both on industry and regulatory side. The real concerns are not clear from this section. It would increase value of the section, if there would be clear articulation of the false conclusions, which one aims to circumvent.	
EUCROF - EU CRO	0	0		EUCROF appreciates the opportunity to provide comments for the Guideline ICH E20 as adaptive designs will become increasingly important in the future, especially in the context of complex clinical trials. As always, more examples would be considered as beneficial, although it is recognized how difficult this is. Maybe some case studies given as examples in an Appendix would be a way to give this Guideline a more practical touch. Other than that, we feel that it is very complete and addresses very important issues.	
EFPIA	234	235	3.51	Trial integrity is introduced to achieve "objectives in a […] timely manner", but the opportunities with adaptive designs with regard to this are not discussed.	Acknowledge the high potential for adaptive designs to accelerate the process of drug development and to allocate resources more efficiently without lowering scientific and regulatory standards.
ACRO					ACRO recommends that E20 could be more forward-looking by acknowledging emerging analytic frameworks such as machine learning-assisted response modeling and real-time data integration. The guideline should consider how future adaptive methodologies can be evaluated under the same principles of pre-specification and statistical rigor, drawing lessons from ICH M15 Guideline on general principles for model-informed drug development.
ACRO					"Ensuring that a prior accurately reflects relevant available information and addressing the potential for conflict between prior and current trial data introduces additional uncertainties that are not present when using frequentist analyses with no borrowing. However, Bayesian analyses are more suitable for evaluating hypotheses, as they directly measure the degree to which data support or undermine hypotheses and probabilities of hypotheses."
BSWG			5.3	A proper characterization of Adaptive Bayesian design is needed. It is already well established in the literature that adaptive Bayesian clinical trial designs refer to a wide variety of clinical trial designs that use Bayesian statistical reasoning and/or calculations in various ways (Berry et al., 2010, FDA Guidance on adaptive design, 2019). For example, a design that applies Bayes methods for interim decisions on stop/go etc. but make inference using frequentist method also fall in this bucket.	
International Advisory Committee on Clinical Trials in Multiple Sclerosis			General	The guideline presents a broadly acceptable framework, grounded in sound principles. Given practical complexities of implementation which are critical to the success of adaptive designs in real-world clinical trial settings, we recommend the inclusion of concrete examples throughout the guideline. These would help illustrate key concepts and provide context for how adaptive design principles are operationalized.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
International Advisory Committee on Clinical Trials in Multiple Sclerosis			General	While we recognize this is a broad guideline, further details and guidance would be helpful. Particularly on such details as 1) requirements for simulation software validation, reproducibility, seed management, and independent code review, 2) minimum safety exposure guidance, 3) evolving estimands when populations or treatments change across stages, and handling intercurrent events with heterogeneous follow-up and 4) recommend including expectations for real-time data quality, validated adaptation algorithms, and integration of decentralized data streams	
International Advisory Committee on Clinical Trials in Multiple Sclerosis			General	This guideline aims to advise "on adaptive designs to accelerate the process of drug development". We believe it may be helpful to also include adaptive designs, such as MAMS-ROCI, which are applicable to many existing/approved agents for which optimisation of dose, frequency, or duration of therapy is desirable (e.g. doi: 10.1016/S1470-2045(23)00095-5).	
International Advisory Committee on Clinical Trials in Multiple Sclerosis			Section 4.3, 5.5	The guideline would benefit from explicit guidance on the integration of properly qualified biomarkers in adaptive trials, particularly for chronic conditions such as multiple sclerosis, where important aspects of the disease process may be subclinical, and where clinical endpoints may take extended periods to manifest. Biomarkers can serve as valuable intermediate outcomes to inform trial adaptations. It is important to acknowledge that surrogate endpoints, such as biomarkers, differ in their validity and predictive value for clinical outcomes, which can influence both trial adaptations and interpretation of results.	
International Advisory Committee on Clinical Trials in Multiple Sclerosis			Section 1	There is a notable absence of discussion on patient-reported outcomes, which serve as primary endpoints in several disease areas, including myasthenia gravis. Their role in adaptive designs warrants specific consideration.	
International Advisory Committee on Clinical Trials in Multiple Sclerosis			Sections 3.5, 5.1	We have concerns regarding the management of interim data access, especially in smaller organizations where structural safeguards may be limited. We recommend the guideline provide clearer direction on who should access interim data and what measures are necessary to preserve scientific and operational integrity.	
International Advisory Committee on Clinical Trials in Multiple Sclerosis			General	The guideline could more clearly delineate the scope of permissible adaptations, particularly those informed by external data sources. This distinction is essential for regulatory clarity and trial planning.	
International Advisory Committee on Clinical Trials in Multiple Sclerosis			Section 3.5, 5.6	Adaptive designs introduce ethical and operational challenges in communicating trial modifications to participants and investigators. The guideline could address best practices for managing information flow when trial arms are altered or discontinued.	
International Advisory Committee on Clinical Trials in Multiple Sclerosis			Section 5.2	This section provides very specific advice (e.g. it may be important to use 100,000 or more repetitions per scenario to ensure sufficient precision for estimating the Type I error probability; and, if custom software is used, to provide the simulation code); however, these expectations may be disproportionate and create unnecessary burden with moderate/low additional regulatory value. Can the requested information be proportional to the complexity and novelty of the adaptive trial design? For example, when a minimal set of scenarios and fewer repetitions would be sufficient. Can the provision of code be optional and only requested if needed, rather than always in advance?	
International Advisory Committee on Clinical Trials in Multiple Sclerosis			Section 5.3	This section is very high level and not very practical; we recommend the addition of specific scenarios where EMA would accept Bayesian adaptive designs for late-phase or registration trials be indicated. It would be beneficial to clarify if real-world data are accepted as source of external information and how the "minimum" amount of concurrent trial data will be established when borrowing external information.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
International Advisory Committee on Clinical Trials in Multiple Sclerosis			Section 4.5	The tone of this section is rather negative, we recommend rephrasing and including an example of an acceptable RAR design.	
International Advisory Committee on Clinical Trials in Multiple Sclerosis			Section 3.5, 5.6	These sections explain that informed consent forms should cover the possibility of adaptive changes and that ethical standards must be maintained when limiting what can be inferred from interim adaptations. We recommend the inclusion of further guidance on patients' consent that is needed for adaptive trials. Particulary to clarify if it is important to re-consent participants after an arm is removed or when safety data are available at the interim analysis. We recommend including guidance as to whether it would be useful to have separate consent at each stage of MAMS trials, since participants should be informed of which arms/stages they are contributing to, while being informed about the overall trial design. Additionally, the guideline could clarify what minimum information about adaptive features should be included in informed consent forms. Perhaps this could be achieved through the inclusion of examples of wording to explain adaptive designs in plain language.	
International Advisory Committee on Clinical Trials in Multiple Sclerosis			Sections 4.2, 6.1	Further guidance is needed on the regulatory and ethical implications of modifying effect size targets mid-trial. Specifically, the process for approval and the impact on trial integrity should be clarified.	
International Advisory Committee on Clinical Trials in Multiple Sclerosis			Section 4.1	Group sequential designs are a form of adaptive design and could be explicitly acknowledged as such within the guideline to ensure consistency in terminology and application.	
Teva Pharmaceuticals	0	0	NA		We appreciate that the current guideline largely focuses on statistical consideration. It would be helpful to expand the scope and recommendations with detailed guidance on other relevant aspects for adaptive designs, e.g., clinical pharmacology consideration, adaptive design for different modalities etc.

2. Specific comments on text

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
ACRO	2	20	1	Challenge: The definition of an "adaptive design" here may be too restrictive by excluding certain data-driven adaptations that are operationally pre-specified but may not meet a strict "prospectively planned" definition. This could unintentionally exclude valid adaptive elements implemented via automated or pre-defined decision rules.	Recommendation: We recommend clarifying that adaptive elements may include structured responses to emerging data trends when pre-defined within the protocol or statistical analysis plan. This approach maintains scientific integrity while reflecting the realities of modern, digitally enabled adaptive trials. We suggest the addition of the following text: "Digitally enabling modern clinical trials may change data distributions in unforeseen ways. Therefore, consideration may be given to pre-specification of adaptive design elements in protocols and/or statistical analysis plans, which allow for data-driven changes to study conduct or analysis."

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EUCROF - EU CRO	2	4	1	"This document provides guidance on confirmatory clinical trials with an adaptive design intended to evaluate a treatment for a given medical condition within the context of its overall development program."	We think it should be mentioned that adaptations are not only possible in confirmatory but also in other settings, e.g., in an early stage of clinical development. Proposed change: Although adaptations are also possible in non-confirmatory settings, this document provides guidance on confirmatory clinical trials with an adaptive design intended to evaluate a treatment for a given medical condition within the context of its overall development program.
Invents consortium - EU Horizon project	2	4	1	This document provides guidance on confirmatory clinical trials with an adaptive design intended to evaluate a treatmnt for a given medical condition within the context of its overall development program.	The ICH E20 guideline provides principles for using adaptive designs in confirmatory clinical trials, but its concepts can also be applied to earlier phases of clinical development.
BSWG	4	6	1	"from participants in the trial" may be too restrictive as there could be accumulating concurrent external data to guide modifications. As long it is recognized and part of the planned modification/adaption, there is no reason why "external data" cannot be used.	
Cancer Research UK Clinical Trials Unit, University of Birmingham	5	6	1	The guidance states that adaptations should only be made due to interims and that those adaptations must be pre-planned, however platform trials which necessitate adaptations often require the addition of new arms. The timing of these additions is often unknown and will impact on the current trial and therefore cannot be prospectively planned and might not be based on accumulating trial data.	If addition of new arms are not deemed to be within scope of the guidance then this should be made clear. If they are in scope the this needs its own section as the preplanned element cannot be relevant and also included in section 4.
EFPIA	6	6	1.01	The term 'interim analysis' has not been defined and given that there is no unique definition in regulatory guidelines of the term, this creates ambiguity. We would suggest using the FDAs Adaptive Design Guideline definition within the E20 guideline, "An interim analysis is any examination of data obtained from subjects in a trial while that trial is ongoing and is not restricted to cases in which there are formal between-group comparisons. The observed data used in the interim analysis can include one or more types, such as baseline data, safety outcome data, pharmacokinetic, pharmacodynamic or other biomarker data, or efficacy outcome data."	Define the term 'interim analysis'
EFPIA	6	8	1.01	Recommend clarifying that such details should be documented not only in the clinical trial protocol but also, where appropriate, in a separate Statistical Analysis Plan (SAP). Specifying adaptation rules and statistical methods in advance in both documents would emphasize their importance and promote regulatory consistency.	Propose change to "*The term prospectively planned means that the potential trial adaptations are pre-specified in the clinical trial protocol (and a separate statistical analysis plan where appropriate) prior to initiation of the trial.* "
EUCROF - EU CRO	6	8	1	The term prospectively planned means that the potential trial adaptations are pre-specified in the clinical trial protocol prior to initiation of the trial.	It is crucial that all potential design adaptations and decision criteria are pre-specified and fully documented in the study protocol and in the Statistical Analysis Plan (SAP).
Invents consortium - EU Horizon project	6	8	1	The term prospectively planned means that the potential trial adaptations are pre-specified in the clinical trial protocol prior to initiation of the trial.	It is crucial that all potential design adaptations and decision criteria are pre-specified and fully documented in the study protocol and Statistical Analysis Plan (SAP).
EFSPI/PSI Regulatory ESIG	7	8	1	"potential trial adaptations are pre-specified in the clinical trial protocol prior to initiation of the trial "	Change to "potential trial adaptations are pre-specified in the clinical trial protocol prior to the first clinical cutoff for an interim analysis".
EFPIA	8	9	1.01	Suggest guidance offers some flexibility for the incorporation of potential trial adaptations even after the initiation of the trial, but still in a pre-specified manner	
Dr.Viviana Mascilongo	10	17		Missing summary of TYPES OF ADAPTATIONS	To add a scheme of the various of Adaptations'Types

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Regeneron Pharmceuticals, Inc.	10	12	1	Regeneron appreciates the Council's efforts in developing guidance to assist sponsors in designing adaptive clinical trials. We respectfully suggest that the scope of this guideline be expanded to include design modifications informed by external sources of information. Adaptations based solely on external data can present similar challenges to other adaptive designs, such as ensuring trial integrity. For example, sample size re-estimation often occurs due to emerging external information, and is currently considered part of adaptive design. Since there are currently few guidelines addressing these types of designs, their inclusion would provide clarity for sponsors.	
EORTC	16	16	1	It is indicated that the guidelines apply to trial addressing risk/benefit. Why should it not apply to later trials addressing effectiveness?	
EORTC	20	20	1	It seems to imply that it applies to all trials? Confirmatory trials? We urge for a clear description of the scope to avoid any ambiguity.	
Teva Pharmaceuticals	20	21	2	"Although the guideline primarily focuses on confirmatory clinical trials, the principles outlined are relevant to all phases of clinical development" • It would be helpful to clarify the statement above. For example, consider clarifying whether the statement implies that the principles can be applied in studies for other development phases.	
c4c-S	21	82	2	clearer reference to empirical evidence and case examples (e.g., oncology, rare diseases)	Add illustrative examples of successful adaptive confirmatory trials and contexts (e.g., platform trials, rare diseases)
ACRO	22	82	2	Challenge: There are several examples provided here, but all are high level and effectively hidden within the text.	Recommendation: More in-depth examples, in a more structured format, would be useful. In Section 2, there are 5 or more uses of "For example" with examples given. These illustrative case examples could be extracted from the main text and expanded with more detailed information to clarify the scenarios discussed. An appendix may be useful for this purpose.
c4c-S	29	38	2	The text is difficult to read as it is.	I recommend to add numbers in the different advantages of the flexibility of adaptative designs, i.e, 1-ethical advantages, 2-improve the efficiency, 3-can help improving uncerstanding …
EFPIA	30	32	2.02	While the guidance mentions the potential for early stopping in a group sequential design as an advantage of an adaptive design, it misses to mention the futility stopping as a key advantage in such designs. A futility stopping rule is a pre-specified modification that uses interim data to determine if the trial should be stopped early due to a low probability of success or some safety signal. Therefore, this is a key advantage in any confirmatory adaptive trials.	Suggest including a sentence in line 32 on the advantage of having futility analysis stating "Second, adaptive designs also allow for early futility stopping according to a pre-specified binding or non-binding method that saves time and money by ending ineffective trials early, protecting patients from treatments with no benefit, and allowing resources to be redirected to more promising research. ". We also suggest including the possibility of achieving accelerated approval process using an adaptive design method as compared to a non-adaptive method.
EFPIA	32	35	2.02	“Second, adaptive designs can improve the efficiency of a trial, for example, by increasing its power for a given expected sample size.” This “for example” is too simplistic as increasing expected sample would be expected as necessary (as in SSR) while maximum SS GSD considering efficacy and futility would be larger than a fixed design any way. So “maximize power for given expected sample size” is a good measurement but a better one would be “maximize (increase) power within the resource allocation limits to the Sponsors”. The resource allocation limits could include the maximal sample size. Reducing expected sample size is a useful metric but often not the primary one, as adaptations may prioritize maximizing power and trial reliability under sponsor constraints, even if expected sample size increases slightly for greater robustness. In many situations, maximal sample size could be too big.	We recommend to replace the phrase “for example, by increasing its power for a given expected sample size” (line 34-35) with 'for example, by increasing its power within available resource constraints' as this better reflects the multifaceted nature of efficiency in adaptive designs.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	32	34	2.02	Sample size limitations are especially rate limiting in rare disease trials. Recommend calling attention to how critical this benefit can be to trials that cannot feasibly enroll a large number of participants.	Recommend adding at the end of the statement "…which can be especially beneficial in rare disease trials where sample size limitations exist."
EFPIA	32	32	2.02	It should be acknowledged here that a key advantage is speed and the ability to get effective medicine to patients that need them faster.	
EFSPI/PSI Regulatory ESIG	33	34	2	"by increasing its power for a given expected sample size" The term "expected sample size" has a fixed meaning in the adaptive design literature to mean the (mathematical) expected value of the sample size of a design, taking into account the probabilities to stop at interim analyses. The formulation here is misleading as it seems that "expected sample size" refers to the actual maximal sample size of the trial.	Remove "expected", as "given" already covers the meaning.
EFPIA	34	35	2.02	Adaptive designs can also help improve understanding of PK/PD relationship.	Third, adaptive designs can help improve understanding of treatment effects*, PK/PD relationship* and decision-making
EFPIA	36	37	2.02	Clarification that this relates to dose selection.	suggest adding the ord of selection in the sentence "…may reduce uncertainty about the dose selection for a better benefit-risk profile…"
EFPIA	39	39	2.03	Adaptive designs bringing complexities should not be shown as a disadvantage as it contradicts the the philosophy of FDA's Complex Innovative Design pilot programme. The focus for any trial design needs to be efficiency in achieving the trial objectives and answering the clinical question of interest. If some complexity in the design methodology enables to achieve such efficiency (eg: accelerated approval) in answering the clinical question as compared to a simple method, such complex methodology needs to be considered as an advantage instead of a disadvantage.	Suggest replacing the word "complexities" with "operational challenges".
EFPIA	39	39	2.03	The list of challenges is appropriate, but refers mostly to not well-designed trials. A proper adaptive design will not implement conventional analysis methods, it will be designed to provide a sufficient safety data base and it will not be implemented, if there is too fast patient enrollment. As such, recommendation to	- Line 39: Replace "However, adaptive designs also present challenges, as they may add complexities and uncertainty related to the key principles" with "If not properly designed and implemented, adaptive designs present challenges, add complexities and uncertainties and related to key principles". - Add to Line 47: That is why conventional analyses will generally not be used. - Add to Line 49: That is why statistical methods are implemented to control the Type I error. - Add after Line 59: Adaptive designs should be planned properly, including operational assumptions and not just statistical considerations.
EFPIA	39	40	2.03	Recommend update "complexities" to "complexity" to ensure parallel structure when listing "complexity and uncertainty". It is also more appropriate because it refers to the general nature of the challenges rather than multiple distinct instances of complexity and uncertainty.	Propose change to "*However, adaptive designs also present challenges, as they may add ~~complexities~~ complexity and uncertainty related to the key principles discussed in Section 3."*

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	39	59	2.03	Highlighting the challenges of analysis of adaptive trials is very welcome. Within the draft guideline two specific topics are highlighted: Type I error control; and valid estimation. Type I error control is an issue because conventional (ie fixed sample) methods of p value computation are not valid for adaptive designs. It would be helpful to highlight this broader point, mentioned in line 952, as well as error control. This is important as "comparing the wrong p value to the wrong threshold to deliver a 2.5% false positive rate" is not an appropriate approach.	Suggest replacing "Type I error control" with "computation of valid p values" in some parts of the document e.g., Line 139 and Line 165. Lines 45-50 would benefit from a short explanation that the increase in Type I error is due to conventional p value methods not being valid. Similar to Line 952, it would be helpful to state specifically that p-value calculations should appropriately account for the adaptive design. Proposed that in section where it is stated that "conventional treatment effect estimates at trial end may be biased", be replaced with "conventional treatment effect estimates at trial end may be biased and the p value may overstate the advantage of the experimental treatment".
c4c-S	39	59	2	The text is difficult to read.	I recommend to add numbers in the different challenges, i.e, 1-add complexities, 2-may require more time, 3-increased Type I error…
EFPIA	40	43	2.03	Confirmatory randomized trials are typically complex—multicenter, multiregional, with centers that differ in clinical practice, language, and other factors—and the planning for such trials, whether adaptive or non-adaptive, requires care to maintain confidentiality and trial integrity, while simultaneously monitoring safety with the associated need for access to unblinded information by safety monitors and IDMCs. In practice, additional logistical complexity associated with the implementation of an adaptive design is relatively minor. The need for access to, and interim analysis of, unblinded information is present in both adaptive and modern non-adaptive trials to ensure ongoing scientific validity and ethical balance.	The guideline should acknowledge the need to consider all aspects of the design, including non-adaptive monitoring requirements, in ensuring appropriate confidentiality and trial integrity.
EFPIA	43	45	2.03	The text mentions the complexity but could be more explicit about the required resources for planning and simulation, which is a major practical consideration for sponsors.	Under the “challenges” paragraph, consider specifying the additional requirements; e.g.: “The planning of a robust adaptive design often requires a greater upfront investment of time and resources for activities such as extensive clinical trial simulations, development of specialized statistical software, and coordination between multiple functional areas, compared to a traditional non-adaptive design.”
IDSWG	47	49	2	It is agreed that the sample size re-estimation based on treatment effect size would lead to Type I error inflation. For the less technical savvy reader. The reason for this inflation should be described	Add the reason why the Type I error rate would be inflated. Also make clear that this sample size re-estimation example is unblinded.
IDSWG	47	50	2	Can we just say increased instead of doubled in the following sentence. Do we know it is exactly doubled? For example, in a design with an interim 47 analysis to modify the target sample size based on the estimated treatment effect, the Type I 48 error probability can be more than doubled when using analysis methods that do not account 49 for the adaptation.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
BSWG	48	48	2	It is important to first clearly define and explain what "Type I error probability" is, as this term has appeared 29 times in the E20 document. Type I error (TIE) refers to the type of wrong decision in which a drug is truly not effective (does not mean its effect is 0 or the same as the control) but the human decision is to conclude that the drug is effective. The chance (probability) of making the TIE is the Type I error probability. That probably can be quantified based on a model, either for Y the data or theta the parameter. For example, if the decision rule is to conclude the drug being effective if Pr(theta > 0 | Y) > 0.95, then the estimated TIE probability is bounded by 0.05, assuming the probability model of computing Pr(theta > 0 | Y) is right. This is a Bayesian quantification of TIE probability. Frequentist computes Pr(|t*| > t_0.025 | theta = 0), or a p-value, and is also the TIE probability for frequentist, where t* and t are observed and theoretical test statistics. It is unclear which TIE probability is being used in this document, and why?	Suggest to define both Bayesian and frequentist versions of Type I error probability. Suggest to allow either to be used for quantifying the chance of making erroneous Conclusions.
EFPIA	48	49	2.03	Lack of context when stating potential doubled type I error.	Suggest either adding the context to describe the scenario of potential doubled type 1 error or modifying the general statement, e.g., ...the Type I error probability can be inflated.
EFPIA	49	49	2.03	The numerical detail ("more than doubled") is too specific and otherwise not appropriate for an ICH guideline that is inherently flexible to allow the regulators to implement.	Propose to replace by "greatly increased"
EFPIA	50	52	2.03	The reporting bias from successful studies also arises in a study without early stopping if only successful studies with T>1.96 are reported. It is unlikely a Sponsor would request a reporting bias correction for the effect estimate of a successful ordinary, single-stage study. To insist on unbiasedness for a conditional estimate of E(Y|Z=z) where z is the observed stopping stage (only if z is small, otherwise not) is inconsistent and not per standard practice.	delete
Cancer Research UK Clinical Trials Unit, University of Birmingham	52	52	2	Explain what 'special' means and offer examples or reword.	Perhaps use the term 'appropriate' instead.
EFPIA	52	52	2.03	An increase in the expected sample size is a major disadvantage of adaptive designs, particularly when initial trial assumptions (eg: initial treatment effect estimate) are overly optimistic or inaccurate. While adaptive designs can provide flexibility, they risk ballooning the study to an unfeasible size, straining resources and potentially exposing more patients to ineffective treatments.	Suggest including this point in line 52 as a potential disadvantage of an adaptive design.
EFPIA	52	53	2.03	The phrase “Therefore, special analysis methods for hypothesis testing and estimation that account for the adaptive design usually need to be used”, in combination with later text, suggests that the usual estimators result in significant bias. While bias often exists, it is also often of insufficient magnitude to represent a meaningful threat to the validity of the conclusions to be drawn from a trial. The Guideline should communicate the importance of quantifying bias, e.g., through simulation, and determining whether the magnitude of bias represents a meaningful threat to the validity of the clinical trial.	We suggest consideration of text to read: “Bias and Type I error under the proposed adaptive design should be quantified, compared to similar non-adaptive approaches and special analysis methods for hypothesis testing and estimation that account for the adaptive design may need to be considered if the bias is found to be substantial.”
EORTC	53	55	2	The potential risk for obtaining less safety information may result in more uncertainty. The guideline should indicate that the use of adaptive design and the risk for less safety information may result in uncertainty in regulatory decision making which should be reflected in the EPAR . This also potentially applies to other uncertainties related to the use of adaptive designs.	
EFPIA	54	54	2.03	CT with adaptive design may provide limited safety data as acknowledged in the document.	In addition to the proposals already included, the document could suggest the Sponsors perform aggregate safety data review at the program level (if available) to inform the B:R assessment at a given single CT with adaptive design

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
c4c-S	60	83	2	This paragraph should have a subtitle.	I recommend to add a subtitle to separate this paragraph to the previous.
EFPIA	61	63	2.04	Clarify the tensions between “confirmatory nature” and adaptive design.	Concerns should be clearly spelled out. From a statistical perspective, appropriate designs fulfilling "key principles" should be acceptable and provide no "tension between confirm and adapt". As such, this subsection is unclear in its objective.
EUCROF - EU CRO	61	63	2	"There can be a tension between the confirmatory nature of a late-stage clinical trial and the proposal to adapt aspects of the trial while it is ongoing."	What is meant with "late-stage"? Late phase? Late timepoint in development but before marketing authorisation? Would be good to be more precise.
ACRO	63	65	2	Challenge: E20 effectively outlines the conceptual advantages of adaptive design but underestimates the practical barriers sponsors face, particularly small or mid-sized organizations, when comparing multiple candidate designs.	Recommendation: We encourage the inclusion of language emphasizing proportionality, ensuring that comparative evaluation should be commensurate with the adaptation’s scale and impact, focusing on rationale for the chosen design rather than exhaustive alternatives. We suggest addition of the following text: “Comparative evaluation should be commensurate with the adaptation’s scale and impact, focusing on the rationale for the chosen design rather than benchmarking against alternatives.”
EFPIA	65	66	2.04	The justification can also include PK/PD considerations.	The justification should include ~~*both*~~ clinical*, PK/PD* and statistical considerations.
EUCROF - EU CRO	65	66	2	"The justification should include both clinical and statistical considerations."	A clinical trial that would produce unreliable results is considered unethical. Adaptations could add uncertainty to produce unreliable results. See the sentence line 66 to 68. Therefore, ehtical considerations should be added. In addition, it should be explained why adaptation is needed. Proposed Change: A clear justification why adaptation is needed should be provided, and should include clinical, ethical and statistical considerations.
Invents consortium - EU Horizon project	65	66	2	The justification should include both clinical and statistical considerations.	A clear justification why adaptation is needed must be provided, and must include both clinical and statistical considerations.
EFPIA	68	78	2.04	CT with adaptive design may provide limited safety data as acknowledged in the document.	In addition to the proposals already included, the document could suggest the Sponsors perform aggregate safety data review at the program level (if available) to inform the B:R assessment at a given single CT with adaptive design
EFPIA	72	72	2.04	A well planned complex design does not affect trial integrity if the complexities are addressing the trial objectives efficiently. A badly planned adaptive process adds considerable uncertainty of maintaining trial integrity. If a design needs complexities to enhance the trial efficiency this needs to be encourage. However, like in a non-adaptive design, even while implementing an adaptive design methodology, it needs to be well planned to acheive the targeted estimand.	Suggest replacing the word "complex" to "inadequately planned".
EFPIA	78	78	2.04	A proposed adaptive design requires a robust justification.	A proposed adaptive design requires a clear*, robust* and compelling justification.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	78	78	2.04	Recommend adjusting the statement: "A proposed adaptive design…". There is a clear statement earlier that efficacy/futility stopping would not add substantial uncertainty, while the challenge is given, when there are multiple trial features which should be adapted. The guidance should be clearer in spelling out, when justification is required, as currently anything requires justification. For example, it is currently not practice to list advantages and limitations for each GSD proposal. This would result in an increase of documentation and additional uncertainty on "acceptable limits" for designs, while the key drivers for acceptability should be given by the named key principles.	Change the statement to "A proposed fixed design without interim analyses...".
EFPIA	78	83	2.04	It is agreeable that any proposed adaptive design requires a clear and compelling justification. However, we argue that this comparison to alternative designs should ensure a fair and equal standard of scrutiny. Specifically, the same level of rigorous justification for key design elements and assumptions should be required of all trials, even when a conventional, non-adaptive trial design is used for them. The existing guidance is well-structured and highlights that the choice of an adaptive design makes the rationale — to address residual uncertainty before a Phase III trial begins — explicit, thereby being upfront about that uncertainty. We believe this same level of candour and rigor should be applied to all trials. Non-adaptive trials should also be required to provide a justification for why an adaptive approach was not used, especially when significant residual uncertainty exists. This would ensure a consistent high standard for all trial designs and promote a more honest and transparent approach to clinical research.	Add at the end of that paragraph a clarification that states "An equally clear and compelling justification is expected to apply to any design (including those with a non-adaptive design) particularly where large residual uncertainty exists that may affect their operating characteristics(e.g., in expected recruitment rates or in expected event counts)."
EFPIA	78	83	2.04	A primary motivation for the use of adaptive designs is the limitations of non-adaptive designs. This text, which states "This justification should discuss how the proposed design addresses inherent needs of the clinical setting and should provide an evaluation of advantages and limitations as compared to alternative designs (including non-adaptive designs), including a comparison of important trial operating characteristics (e.g., power, expected sample size, reliability of adaptation decisions) between candidate designs." should be modified to explicitly include the specific limitations of the non-adaptive design(s) that the adaptive design mitigates or addresses, e.g., quantitative assessments of limitations in power or the in the selection of an optimal dose associated with a non-adaptive approach. The motivation for the adaptive design is inadequately communicated without identifying the limitations of a non-adaptive design in the proposed setting.	The cited text should be modified to explicitly include the specific limitations of the non-adaptive design(s) that the adaptive design mitigates or addresses, e.g., quantitative assessments of limitations in power or the in the selection of an optimal dose associated with a non-adaptive approach.
Regeneron Pharmceuticals, Inc.	78	83	2	Regeneron agrees that sponsors should provide a clear and compelling justification for a proposed adaptive design. We also note that the expectations regarding evaluations of alternative designs may create an unnecessary burden for sponsors. Additionally, in some cases, the adaptation may be motivated by clinical considerations where comparative operating characteristics may not meaningfully reflect the rationale for the design. For these reasons, Regeneron proposes removing certain wording from the guideline, as proposed in Column G.	ORIGINAL TEXT: This justification should discuss how the proposed design addresses inherent needs of the clinical setting and should provide an evaluation of advantages and limitations as compared to alternative designs (including non-adaptive designs), including a comparison of important trial operating characteristics (e.g., power, expected sample size, reliability of adaptation decisions between candidate designs. PROPOSED TEXT: This justification should discuss how the proposed design addresses inherent needs of the clinical setting and should provide an evaluation of advantages and limitations as compared to alternative designs (including non-adaptive designs), including a comparison of important trial operating characteristics (e.g., power, expected sample size, reliability of adaptation decisions between candidate designs.
EFPIA	79	83	2.04	Please clarify how the reliability of adaptation decision is an operating characteristic that can be quantitatively assessed as a basis for comparing diffferent designs. The way it's currently written is not clear.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
ACRO	84	307	3	Challenge: The five foundational principles are well-structured but would benefit from stronger operational linkage to existing ICH guidelines.	Recommendation: Specifically, Principle 3.5: "Maintenance of Trial Integrity" should explicitly reference the digital controls, audit trails, and system validation practices detailed in ICH E6(R3) and E8(R1). Reinforcing these cross-references will ensure that adaptive trials uphold integrity and transparency across digital data systems.
c4c-S	84	166	3.1–3.2	clarify how regulators will assess adequacy of simulations	Recommend inclusion of practical expectations (e.g., minimum number of simulated scenarios, Type I error precision) and reference to Appendix/technical note
EFPIA	86	86	3.01	Adaptive designs are used to not only ensure reliability and interpretability of results but are also used to enhance trial efficiency.	Suggest including "and enhance trial efficiency " at the end of the sentence in that line.
EFSPI/PSI Regulatory ESIG	86	0	3	In several instances in the guidance the words "reliable", "interpretable", "validity" appear. However, these terms are nowhere defined, or at least some context is given for what they are supposed to refer to.	Remove these terms, or provide guidance on what is meant by them.
Cancer Research UK Clinical Trials Unit, University of Birmingham	89	90	3	The guidance states that the principles should be followed regardless of whether frequentist or Bayesian, however it is written from a frequentist stance. Should the terminology throughout the document be appropriate for both approaches. Is the guidance intending that all trials report the type I error irrespective of design as this is inappropriate for Bayesian designs.	Updates throughout to clarify position or be clear in this section. Specify whether frequentist parameters are required to be reported even if the trial framework is Bayesian.
EORTC	93	93	3	It could be important to indicate that development is not the end of the process. Trials designed for access, such as optimisation trials should also be considered. Guidelines are important document to change the culture of drug development into access as many questions do remain beyond licensing.	
c4c-S	100	100		Suggest adding explicit reference to "intended populations, including paediatric subpopulations where appropriate," as adaptive designs may be essential in early paediatric proof-of-concept or bridging trials, where the evidence base is often limited and recruitment constraints are significant.	
EFPIA	101	102	3.12	There is an opportunity to improve clarity here, in the sentence that reads "The number and complexity of adaptations at the confirmatory stage should generally be limited." As it stands, this statement is overly broad. What is overly complex and what number of adaptations is too many is a function of the particular clinical setting, the details of the proposed trial design, and the needs of the development program. In lines 108-118 the guideline provides an example design that is, in fact, a good design in one context and perceived to be inadequate in another. Determining whether a design is too complex or requires too many interim analyses is highly dependent on the context.	Please revise to state: "The complexity of proposed adaptations and the number of interim analyses should be thoroughly investigated and appropriate for the context of the trial." Otherwise, the Guideline risks being used to support a contention that a proposed trial design cannot be considered confirmatory because it has too many interims, a position that is unsupportable out of context.
c4c-S	101	107	3.1	The guideline discourages complex adaptive confirmatory designs and overemphasizes limiting the number of adaptations. This may unnecessarily constrain innovative, statistically valid adaptive methodologies supported by extensive simulations.	Clarify that complex adaptive designs (e.g., multi-arm multi-stage, response-adaptive randomization) may be acceptable provided that simulation-based validation demonstrates Type I error control, interpretability, and maintenance of trial integrity.
c4c-S	101	107	3.1	The current text seems to be an adequate compromise as it allows adaptations, but discourages having too many.	No change

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EUCROF - EU CRO	102	102	3.1	"Increasing either of them, …"	**Proposed Change:** Increasing the number and/or the complexity of adaptations, …" Would ease the readability
Teva Pharmaceuticals	103	104	4.1	"The number and complexity of adaptations at the confirmatory stage should generally be limited" • Further clarity is requested with the concept in the statment above. It would be helpful to cearly define the restriction and condition related to this 'number'. Alternatevly, it would be helpful to specify if this would be determined on a case-by-case basis.	
EFPIA	104	107	3.12	A confirmatory trial with multiple adaptations should decrease the number of exploratory trials to be conducted.	Before planning a confirmatory trial with multiple adaptations, sponsors should discuss whether additional exploratory trials are necessary to investigate the question(s) addressed by the proposed adaptation(s). A confirmatory trial with multiple adaptations should decrease the number of exploratory trials to be conducted.
EFPIA	105	105	3.12	If the trial can establish superiority at control of type-1 error with use of multiple adaptations, then the rational for conducting additional exploratory trials is unclear. At time of initiating a confirmatory trial, there is generally sufficient understanding and remaining uncertainties are expected to be low. The sentence is not adding relevant content to the guidance, as any intervention owner will carefully assess, whether an additional exploratory trial would significantly change perspectives vs. current knowledge. The likelihood that the exploratory trial would result in significant changes, which would not have been recognised in the confirmatory trial, is unclear.	Suggest removing lines 104-107.
EFPIA	108	118	3.13	A seamless phase 2/3 study, integrating dose finding with evaluation of possibly more than one dose in the second stage, may well have better operating characteristics and probability of selecting the optimal dose than the combination of a traditional dose-ranging study followed by the evaluation of a single dose in phase 3. The implication that a seamless design is generally inferior is overly broad and often incorrect, as the relative performance of the two strategies depends on the details of each trial design. For example, a seamless 2/3 design may naturally enroll more patients in dose finding, with improved dose selection. Further, the statement that "(a)n adaptive design should generally not serve as a replacement for a proper dose-ranging trial" implies that an adaptive trial cannot perform equivalently or better than a traditional dose-ranging trial, which is not true; again, the relative performance depends on the specifics of each trial's design.	This section should likely be removed, and this can be done without loss of continuity of the document. The statement that "(a)n adaptive design should generally not serve as a replacement for a proper dose-ranging trial" implies that an adaptive trial cannot perform equivalently or better than a traditional dose-ranging trial, which is not true; again, the relative performance depends on the specifics of each trial's design. This statement should be modified to acknowledge that the better approach depends on the specifics of each trial design.
Cancer Research UK Clinical Trials Unit, University of Birmingham	115	118	3.1	For seamless trials across two or more phases these adaptations could encompass a dose ranging component. Is the guidance implying that this would not be considered confirmatory. If not then this needs to be clearer in the text.	
EFPIA	115	118	3.13	Flexibility should be included in how this is laid out with respect to sequential drug development which may not be always possible in rare disease or difficult to recruit patients	
EFSPI/PSI Regulatory ESIG	115	116	3.1	"An adaptive design should generally not serve as a replacement for a proper dose-ranging trial" "proper" dose ranging trial - this could be interpreted to suggest that a trial with adaptive design is not a proper trial.	Replace "proper" with "dedicated".

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	117	117	3.13	The document only uses the term "doses". The more general term "dosing regimen" (consisting of dose, dosing frequency, loading/maintenance dose usage, ...) describes drug development realities better.	Replace "dose" by "dosing regimen" throughout the document
Breakthrough T1D	120	121	3.2	Given the growing emphasis on patient input, the guidance should address the conditions under which patient-reported outcomes (PROs) may be incorporated into endpoints or composite endpoints for interim analyses. Specifically, it would be helpful for the guidance to clarify how PROs could be used to support pre-specified decisions, such as early stopping for futility or success. This would enable developers to design trials that are better reflective of patient experience. The EMA should request ICH include more discussion of appropriate selection of endpoints including PROs and biomarkers.	
c4c-S	120	127	3.2	The current text does not require standardized documentation or validation of simulation studies, which are critical for assessing operating characteristics.	Add explicit requirement: “A comprehensive simulation plan and report should be submitted, including input assumptions, seed control, validation checks, and summary of operating characteristics (power, Type I error, probability of correct selection).”
c4c-S	120	127	3.2	The current text does not require standardized documentation or validation of simulation studies, which are critical for assessing operational characteristics (agree with Maria (c4c-S)'s remark above)	See Maria (c4c-S)'s text above
EFPIA	124	126	3.21	Per comment eariler, flexibility should be included on absolute need for these features to be included before initiation	
Teva Pharmaceuticals	124	127	4.2	"If a confirmatory clinical trial is planned with an adaptive design, the number and complexity of adaptations should generally be limited and there should be a justification for adapting aspects of the trial at this stage of drug development" • Further clarity is requested with the concept in the statment above. It would be helpful to cearly define the restriction and condition related to this 'number'. Alternatevly, it would be helpful to specify if this would be determined on a case-by-case basis.	
c4c-S	125	131	3.2	The text is difficult to read.	I recommend to add a list with the specified aspects to take into account prior to initiate a trial with an adaptative design.
EFPIA	127	127	3.21	The phrase “…anticipated rule governing the adaptation decision…” suggests that a precise rule is not required in an adaptive design, which is inconsistent with other sections of the Guideline and with requirements for well-defined operating characteristics. For example, we would not say the use of 1:1 randomization is an “anticipated” feature of a simple non-adaptive trial, and a change from this approach would be considered a major deviation from the intended design. Similarly, if a prespecified rule is not followed, e.g., due to safety data that were not a prespecified component of a dose selection rule, then that represents a deviation from the original design. The word “anticipated” appears 25 times in the draft Guideline, mostly in text that implies that prespecified rules are to be considered flexibly and in a non-binding manner, at least in the sense that modifications of these rules are allowed “within” the adaptive trial design. An IDMC is charged with safety and trial integrity, and thus must be free to recommend changes when prespecified rules clash with those charges. However, the IDMC must include in this assessment the threat to trial validity resulting from any change that invalidates the trial’s operating characteristics.	Here and elsewhere, the Guideline should emphasize the critical role an IDMC has in determining when a deviation from the prespecified adaptive design is appropriate without implying that the IDMC should have a role determining how and when an adaptive rule is applied as an integral part of the prespecified design.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
IDSWG	131	131	3.2	Clarify where all candidate estimands should be specified	"all candidate estimands should be fully pre-specified in the clinical study protocol (CSP) and statistical analysis plan (SAP)"
ACRO	132	137	3.2	Challenge: Consideration should also be given to any digital technology used in the trial and the capabilities of the system to 'adapt' as needed in accordance with the trial. Modifications to technology can take time and may not be able to be built in advance so careful consideration must be given as to how new or different features are turned on/off and released along with the adaptive trial schedule.	Recommendation: ACRO recommends addition of the following statement to ensure the technological aspects are appropriately considered: "Use of any digital technology must be appropriately planned for. This should include consideration of development and implementation of any modifications to the digital technology within the adaptive trial schedule."
EFPIA	132	132	3.22	"Some types of adaptive designs may require more planning than others." is a commonplace.	delete
Invents consortium - EU Horizon project	138	139	3.2	Adequate planning facilitates the evaluation of the appropriateness of the statistical approach for many types of adaptations.	The guideline does not discuss the use of specific statistical methods.
EFPIA	139	141	3.23	The sentence starting with "For example, Type I error probability …..". This is true for group sequential designs. However, Type I error rates needs to be also controlled for blinded sample size re-estimation designs in non-inferiority and equivalence trials (for example : while handling biosimilars) and do not require the pre-specification criteria for early efficacy stopping rules.	Suggest adding "for group sequential designs" after "Type I error probability control".
EFPIA	143	144	3.23	This is the first time an IDMC is mentioned and it seems the guideline is suggesting only an IDMC should make recommendations in adaptive trials (see line 257). In the draft FDA DMC guidance there is mention of a separate independent adaptation body. In terms of trial integrity this approach may add additional complications but is a second committee/body reviewing data not supported at all by this draft guideline?	Provide guidance on whether adaptation committees can add value in a clinical trial with an adaption
EFPIA	143	148	3.23	The statement that adequate planning and pretrial discussion "…ensures the IDMC is prepared to review interim results and make adaptation recommendations during the trial while also protecting individual trial participants' safety." further conveys the concept that the IDMC decides, based on information not captured by the prespecified adaptive rule, whether or not specific adaptations are implemented.	This section be modified to read "…ensures the IDMC is prepared to review interim results and confirm that the application of the prespecified adaptive rules remains scientifically and ethically appropriate, and that no deviation from the prespecified design is necessary to protect individual trial participants' safety." This allows the pivotal role the IDMC plays in maintaining scientific appropriateness and protection of participants to be emphasized, without undermining the importance of prespecification of the adaptive rules in maintaining the designed operating characteristics.
c4c-S	145	148	3.2 & 5	IDMC responsibilities for adaptation oversight are described qualitatively but lack clarity on required documentation and independence assurance.	Specify that: "IDMC charters should detail decision authority, data access procedures, and independence safeguards. The composition, expertise, and independence statement should accompany regulatory submissions." Establish IDMC with expertise in adaptive designs; include statisticians for interim review. - Define IDMC responsibilities, meeting schedules, confidentiality rules, reporting to sponsor.
EFPIA	146	148	3.23	Add a reference to Section 5.1, since the IDMC is discussed more in depth there.	Add a reference to Section 5.1, since the IDMC is discussed more in depth there.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	149	165	3.24	This section acknowledges benefit-risk may justify deviating from a pre-specified rule but still requiers pre-specifying factors that might justify deviation. It is not possible to realistically simulate and control for every "what if" - external events like competitor approvals, changes in standard of care, safety findings from another like-mechanism program, etc. The Adaptiveness loses value if over-constrained thus recommend to also permit for more generalized pre-specified "principles" (e.g. permitting recalibration of statistical threshold if external events change thinking or observed variability differs materially from assumptions). This approach would require justification and documentation of deviations and make appropriate statistical adjustments (alpha reallocation, sensitivity analysis, etc)	Recommend adding clarification that the adaptation rule does not apply to changes made in response to external data.
EFPIA	150	154	3.24	The text: "The extent to which the anticipated rule governing the adaptation decision needs to be adhered to at an interim analysis, however, can vary depending on the type of adaptation and the statistical inferential methods being used. It is generally recommended to use analysis methods that provide valid inference while allowing flexibility to deviate from the anticipated adaptation rule based on the overall benefit-risk assessment at an interim analysis." may be interpreted to suggest that highly conservative analysis methods that maintain type I error control over a wide range of possible interim decisions be used, without addressing the attendant loss of statistical efficiency and potential ethical implications (e.g., requiring a larger number of participants than would otherwise be required). The choice regarding an analysis method that allows this flexibility without compromising type I error control should be based on quantitative or semiquantitative assessment of the relative risks of the two approaches, rather than one approach being "generally" preferred.	The text should be modified to clarify the two competing considerations, namely (1) the ability to maintain valid inference despite flexibility in adherence to prespecified decision rules; and (2) the ability to optimize operating characteristics conditional on the assumption the prespecified rule will be followed. Since both goals cannot, in general, be realized simultaneously, the choice regarding the analysis method should be based on quantitative or semiquantitative assessment of the relative risks of the two approaches, rather than one approach being "generally" preferred.
c4c-S	151	163	3.2	The draft encourages flexibility in deviation from adaptation rules but offers no clear criteria or statistical guidance to ensure valid inference.	Add clarification: "Any deviation from pre-specified adaptation rules should be documented with rationale, and accompanied by sensitivity analyses to demonstrate robustness of Type I error control and inference validity."
EFPIA	152	154	3.24	Simliar statements appear three times in Section 3.2, 4.3, and 4.4. They can be consolidated into Section 3.2. For example, the statement in Section 3.2 refers population selection and treatment selection as examples. Or section 4.3 and 4.4 can have their own examples of flexibilities. Section 3.2 (Line from 152 to 154) It is generally recommended to use analysis methods that provide valid inference while allowing flexibility to deviate from the anticipated adaptation rule based on the overall benefit-risk assessment at an interim analysis. Section 4.3 (Line from 465 to 468) Methods are generally recommended that allow flexibility in deviating from the anticipated adaptation rule, as considering the totality of information available at the interim analysis helps ensure appropriate population selection. Section 4.4 (Line from 505 to 509) It is generally recommended to use methods that allow for flexibility in deviating from the anticipated adaptation rule. Such flexibility enables consideration of the full scope of information available at the interim analysis, helping to support more informed and appropriate treatment selection decisions.	
EFPIA	153	153	3.24	Term "valid inference" is not defined.	Be explicit about what "valid inference" implies

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	163	163	3.24	The guideline recommends to "outline factors that may lead to such deviations", in a section that precisely acknowledges that there may be unexpected events, and emphasizes the need to allow flexibility. Besides the obvious factors (such as safety concern, which would often be included in the decision rules), it is difficult (if even possible) to determine in advance what could be the unexpected concerns that may lead to deviation from the pre-specified rule.	Remove the "and outline factors that may lead to such deviations", "When using statistical methods that allow for the flexibility to incorporate such benefit-risk considerations at the interim analysis, the pre-specified plan should acknowledge the possibility of deviations from the rule. and outline factors that may lead to such deviations". Or acknowledge that some non-anticipated factors may also lead to deviations. "When using statistical methods that allow for the flexibility to incorporate such benefit-risk considerations at the interim analysis, the pre-specified plan should acknowledge the possibility of deviations from the rule and outline factors that may lead to such deviations, however non-anticipated deviations may be justified on a case by case basis" .
EFPIA	163	166	3.24	Suggest broadening this to include the statistical analysis plan (SAP) in addition to the protocol. While the protocol typically contains the core information, the SAP is the appropriate place to detail the specific rules, assumptions, and adherence requirements related to statistical methods. Explicitly acknowledging the SAP would better reflect common practice.	Propose change to "If the planned statistical methods instead require strict adherence to the rule governing the interim decision to ensure valid inference (e.g., Type I error probability control), the importance of adhering to the rule should be documented in the trial protocol and statistical analysis plan."
EFPIA	163	166	3.24	The phrase “If the planned statistical methods instead require strict adherence to the rule governing the interim decision to ensure valid inference (e.g., Type I error probability control), the importance of adhering to the rule should be documented in the trial protocol.” is critically important but could be misconstrued as suggesting this is not usually the case when, for a confirmatory trial, it is generally true.	Phrasing should be added to clarify that strict adherence to the rule is generally required to achieve the specific design operating characteristics.
c4c-S	167	185	3.3	Frequentist Type I error control is clearly emphasized, but corresponding Bayesian criteria are deferred to Section 5.3 without concise principles here.	Include a brief cross-reference: “For Bayesian adaptive designs, equivalent control of false-positive conclusions should be demonstrated via posterior error probability thresholds or operating characteristic simulations, as described in Section 5.3.”
c4c-S	167	185	3.3	The current text, with a reference to Bayesian methodology (section 5.3), is sufficient for me	No change
EFPIA	168	174	3.31	There is no mention in this section on “erroneous conclusions” about type II error. This should also be discussed, as such errors also compromise population heath, and also to be consistent with existing text on lines 196-198.	New material should be added addressing the population health and societal costs associated with type II errors, in the context of the broader regulatory responsibility to optimize population health.
IDSWG	169	174	3.3	An essential element of regulatory decision making is not only controlling false positive efficacy conclusions but false negative conclusions. In any adaptive decision ruiles both false positive and false negative rate need to be controlled	"An essential element of regulatory decision-making is controllng the chances of false positive and false negative conclusions (i.e., conclusions that truly inefficacious treatments or potentially efficacious treatments are considered inefficacious at the time of the interim decision rule). The common approach is to limit the probability of false positive and false negative efficacy conclusions within a trial by using frequentist methods that control the Type I and Type II error probability for a hypothesis test of the primary estimand at a pre-specified threshold (ICH E9).

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
BSWG	171	202	3.3	It is necessary to explain why the default analysis/design should be frequentist for a clinical trial. Is it due to tradition or is it based on a scientific reasoning? The term Type I error probability is used as the primary criterion in assessing the chance of erroneous conclusions. How should it be computed? For example, for most drugs that come to phase 3 trials, their treatment effects are highly unlikely to be zero (or the same as control) or they would not have reached this far. Therefore, a more realistic quantification of Type I error might be false positive rate, in which the drug may assume to have none zero but not clinical meaningful therapeutic effects.	Suggest to change the text on line 185 to "Although the predominant rpparoches to the design and analysis of clinical trials have been based on frequentist statistical methods, they are largely based on tradition rather than demonstrated superiority over alternative methods. Bayesian approaches may also be considered as long as they can clearly demonstrate different types of error probabilities may be controled."
EFPIA	171	174	3.31	While the guideline outlines one "common" approach to control false positive conclusions, it does not mention any alternative approaches. So it remains unclear what alternative approaches can be used to control false positive conclusions (beyond the common frequentist approach). This creates ambiguity and uncertainty.	Rewrite the paragraph making clear that control of the frequentist trial-specific type 1 error control is required (unless an alternative approach is also possible).
Regeneron Pharmceuticals, Inc.	171	174	3.3	Regeneron recommends clarifying that the type I error probability is typically controlled at the clinical trial level as the family-wise error rate rather than focusing on a single hypothesis. To ensure clarity and consistency between guidelines (E9 vs. E20), Regeneron proposes updating the guideline wording in this draft as proposed in Column G.	ORIGINAL TEXT: The common approach is to limit the probability of false positive efficacy conclusions within a trial by using frequentist methods that control the Type I error probability for a hypothesis test of the primary estimand at a pre-specified threshold (ICH E9). PROPOSED TEXT: The common approach is to limit the probability of false positive efficacy conclusions within a trial by using frequentist methods that control the overall probability of Type I error (ICH E9).
EFPIA	175	176	3.32	The paragraphs states that for 'most adaptive designs' it is necessary to control the type 1 error and then provide an example where it is needed. Which is a quite classical case. An example of where there could be flexibility would be of great value to better understand the paragraph	
IDSWG	175	180	3.3	It is agreed that special methods are needed to control the Type I error probability. In addition one must be able to show that your methods result in a reasonable conditional power as well to control Type II errors. One also should mention here how this would be handled in a Bayesian framework	Add text to also emphasize the importance of controlling Type II error rate (conditional and unconditional power) as well as Average Power (In Bayesian context).
IDSWG	177	180	3.2	Can this sentence be simplified? It is not clear. When an adaptive trial design includes multiple testing approaches to control the Type I error probability across multiple primary and/or secondary endpoints, those approaches should additionally address the potential for an increased Type I error probability due to the proposed adaptation.	
Invents consortium - EU Horizon project	177	180	3.3	When an adaptive trial design includes multiple testing approaches to control the Type I error probability across multiple primary and/or secondary endpoints, those approaches should additionally address the potential for an increased Type I error probability due to the proposed adaptation.	(see EMA multiplicity guideline CPMP/EWP/908/99).
EFPIA	181	185	3.33	This discussion on frequentist vs. Bayesian feels out of place in this section and removing it would improve the logical flow of the draft Guideline.	Please consider removing this discussion.
IDSWG	181	185	3.3	Agree on the importance of Bayesian methods should be mentioned earlier in this section	Move this paragraph to the beginning of the section.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	186	198	3.34	This section erroneously implies that an adaptive design results in less data or information. For a trial with early stopping or sample size re-estimation, compared to the trial with full follow-up or a trial that goes to maximum N, the adaptive trial may result in less data. However, it is often true that if the adaptive trial were not selected, the sponsor would have chosen a smaller or shorter-duration trial. Thus, the inclusion of adaptations can often be accurately reframed as "going longer" or increasing sample size—but only when necessary—relative to the non-adaptive trial that would have been conducted. Adaptive trials with multiple doses often result in more data on the target dose. Additionally, this section focuses on safety and secondary endpoints. It is important in general to ensure adequate power for all endpoints and limit type II error.	This section should be expanded to discuss type II error in general and acknowledge that a well-designed adaptive trial may result in greater information, precision, and power in evaluating outcomes associated with the arm(s) that ultimately turn out to be of greatest regulatory importance.
c4c-S	187	194	3.3	While the text stresses efficacy, it insufficiently addresses the risk that adaptive designs (especially early stopping) may yield incomplete safety data.	Recommend adding: "Adaptive designs should ensure sufficient exposure and follow-up for safety evaluation. Simulations should assess safety data sufficiency under potential early-stopping scenarios."
c4c-S	187	194	3.3	It is indeed important that potential side effects are picked up in a trial. Current text may be expanded	See Maria (c4c-S)'s text above
EFPIA	189	192	3.34	"For example, when planning a trial with the potential to stop early for an efficacy conclusion, it is important to justify that the sample size and duration of follow-up at an interim analysis can adequately support a reliable benefit-risk assessment." The benefit–risk assessment is typically based on the totality of data available across the program, rather than on a single study, unless the NDA submission is supported by only one fairly large trial.	We suggest either adding such a condition or stipulating it within the benefit–risk assessment conducted by the eDMC at the time of the interim analysis.
EFPIA	196	197	3.34	This statment is very vauge - "Finally, adaptations can impact the chance of a false negative efficacy conclusion (i.e., lack of evidence of an effect for a truly efficacious treatment) such that it is important to evaluate whether the trial achieves adequate power".	Please include recommendations on how to do this and touch on unconditional power (study level) vs totality of development program power (e.g., combined ph2/ph3 power including power of taking an efficacious dose instead of including dropped doses in Phase 2 which will dilute study power). Clarify if it is sufficient to present conditional power of dose selected and overall program power versus unconditional power in one trial which includes all doses.
EFPIA	196	198	3.34	Please see prior comment re Lines 168-174.	Please see comment above, re Lines 168-174.
IDSWG	196	198	3.3	This paragraph can be integrated into the revisions provided when discussion Type I error control	See revisions above to lines 169-174.
BSWG	200	210	3.4	The guidance indicates that reliability of safety outcomes should be evaluated. In adaptive designs where safety outcomes are correlated with the efficacy endpoint(s) driving adaptation decisions (e.g., dose selection, sample size re-estimation), these safety results may exhibit similar bias to that observed in efficacy outcomes. The current guidance does not clearly specify whether bias correction methods analogous to those required for efficacy assessment must be applied to safety analyses. Please add text clarifying: (1) when specialized analytical methods are necessary for safety assessment, and (2) what level of rigor is expected when correlation between safety and efficacy endpoints exists. On this basis, the guidance may discuss the benefit and risks based on efficacy and safety more.	
BSWG	200	210	3.4	For adaptive designs lacking established bias correction methods for point and interval estimation, the guidance should explicitly state whether investigators are expected to: (1) quantify and acknowledge the bias through simulation studies, or (2) develop and implement bias correction methods. Clear direction should be provided regarding the minimum requirements when validated estimation methods are not yet available for a specific adaptive design.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	200	232	3.4	This section gives example of group sequential design methods without covariate adjustments. However, reliability of estimation specificaly suffers in such designs when covariate-adjustments are implemented and the covariate-adjusted estimators in such cases may lack the asymptotic independent increments structure that is required to directly apply standard stopping boundaries for group sequential designs.	A mention of this challenge in the estimation of the treatment effect for group sequential designs is suggested to be included as an example in this section and appropriate measures such as linear transformation to the sequence of adjusted estimators across analysis times can be suggested as a remedial mesure to this problem.
c4c-S	200	221	3.4	The draft highlights bias and variance considerations but does not specify acceptable levels or reporting expectations for estimation bias or CI coverage.	Add guidance: "Sponsors should report bias magnitude, coverage probability, and mean squared error metrics from simulation studies. Bias greater than 5% of the true effect estimate should trigger justification and mitigation measures."
EFPIA	201	203	3.41	It appears that the evaluation of reliability of safety outcomes is required. When safety outcomes are correlated with efficacy endpoint(s) used to guide study design adaptations, the safety results may also be subject to bias. It is unclear whether specific statistical methods -similar to those applied in the efficacy assessment - should be emplyed for the safety evaluation.	Suggest to include further dicsussion on the safety assessment to justify the need, or lack thereof, for applying specific statistical methods.
Teva Pharmaceuticals	203	236	4.4	• Please consider that controlling Type-1 error, sample size and timing of IA can remain the same issue in non-adaptive trial studies. Accordingly, it would be helpful to clarify whether the sponsor should demonstrate superiority of propose of the adaptive design over other options e.g. by simulation.	
EFPIA	205	208	3.4	The section could explicitly mention that bias can also affect the estimation of safety parameters, not just efficacy, which can be equally important for the overall evaluation of the drug.	Consider adding a sentence: "Sponsors should also consider the potential for biased estimation of key safety outcomes, as this could impact the interpretation of the overall benefit-risk profile."
EUCROF - EU CRO	205	208	3.4	"Sponsors should evaluate bias and variability of treatment effect estimates, including measures such as the mean squared error. In the trade off between bias and variance, the expectation is generally for limited to no bias in the primary estimate of the treatment effect."	The discussion of potential bias appears to be focused entirely on frequentist methods and not on other methods, e.g., bayesian approach.
Invents consortium - EU Horizon project	205	208	3.4	Sponsors should evaluate bias and variability of treatment effect estimates, including measures such as the mean squared error. In the trade off between bias and variance, the expectation is generally for limited to no bias in the primary estimate of the treatment effect.	The discussion of potential bias appears to be focused entirely on frequentist methods.
EFPIA	206	206	3.41	The Mean Square Error is brought forward as measure that needs to be evaluated. This is not measure often reported in clinical trials and will be difficult to interpret. It would be better to report bias and variance separetely which are more easy to interpret and evaluate	
EFPIA	206	208	3.41	The guideline mentions the bias-variance trade-off, with a clear preference for bias, as if it was not a trade-off. However, It is well known that some methods such as the uniformly minimum variance conditional unbiased estimate (UMVCUE) can reduce bias at the cost of a substantial (and often unacceptable) increase in variance, whereas simple methods such as maximum likelihood estimates have limited bias.	Remove the sentence "~~In the trade-off between bias and variance, the expectation is generally for limited to no bias in the primary estimate of the treatment effect~~".
EFPIA	206	208	3.41	The text "In the trade-off between bias and variance, the expectation is generally for limited to no bias in the primary estimate of the treatment effect." suggests that bias should be minimized without respect to the effect on variance and without any quantification of the relative contributions of bias and variance in the validity of the treatment effect estimate or the probability of the trial reaching the correct overall conclusion regarding treatment efficacy.	The text should be modified to mention that higher variance with the associated type II error may be a much greater risk than the effect of bias on type I error risk and, only by quantifying the relative risks, can an optimal design be determined.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFSPI/PSI Regulatory ESIG	208	210	3.4	"The primary analysis should also support calculation of accurate measures of uncertainty such as confidence intervals with targeted coverage probabilities." This might be methodologically possible for some designs, but not for others. What is the implication then? Does this make complex designs, for which inference may not be entirely clear, inacceptable? Or would valid inference based on simulations (e.g. determining maximal potential bias through a simulation study for a broad range of scenarios) be acceptable?	Clarify whether this is mandatory, and what to do if methods are not available.
Regeneron Pharmceuticals, Inc.	208	210	3.4	There are statistical and methodological challenges in estimating treatment effects without bias in certain adaptive designs. The guideline should provide recommendations for situations where no methodology exists to obtain reliable estimate (e.g., confidence intervals with targeted coverage probabilities). It would also be helpful to clarify whether conducting simulation studies to assess potential bias is an acceptable means that can be used to assess such situations.	
EFPIA	209	209	3.41	Confidence Intervals with targeted coverage are brought forward. At times these will not be well-understood and depending on estimation method might not be meaningful. A suggestion would be as with the mean square error to be less direct on specific measures to be reported but more on important concepts to consider which are independent of the statistical estimator	
EFPIA	209	210	3.41	In addition to achieving the targeted coverage, confidence intervals should be informative to ensure their interpretability and relevance. Confidence intervals are informative if they provide more information than the mere hypothesis testing decision. See for example Robertson et al.: "For example, consider testing the null hypothesis H0: $\theta = \theta 0$ vs the alternative H1: $\theta > \theta 0$. If H0 is rejected, then a CI is only informative if $L(X) > \theta 0$", where L(X) is the lower confidence limit. Robertson et al. Confidence Intervals for Adaptive Trial Designs I: A Methodological Review. Stat Med 2025.	…confidence intervals with targeted coverage probabilities. The use of informative confidence intervals, providing more information that the hypothesis testing decision, is recommended.
EFPIA	214	225	3.42	For your example of selecting a high-dose at an interim look leading to an inflated treatment effect estimate, is there a reference for this effect? Similarly, reference is made on Line 222 of group sequential methods, as an example, that may help reduce bias.	The provided examples are very insightful and helpfully illustrate the points being made. Including references for these examples would further enhance their value.
IDSWG	216	217	3.4	The following sentence is not clear to me. Doesn't one adjust for type I error for selecting one dose from several doses at the interim? This holds true even if selection is based on an endpoint expected to be predictive of efficacy rather than the primary endpoint itself.	
EFPIA	217	218	3.42	Secondary endpoints are briefly touched upon and guidance as to the reporting of these would be greatly appreciated	
EFSPI/PSI Regulatory ESIG	217	218	3.4	The statement "Similarly, treatment effect estimates for secondary endpoints may be biased in the presence of adaptations" seems to imply that "reliable" (bias-adjusted/-reduced and accurate) estimates should be provided across ALL efficacy endpoints that are consider relevant for the label. In lines 331 - 334 in Section 4.1 there is an emphasis on the 'primary' treatment effect estimate and associated CI that should adjust for the IA nature. Could it be clarified why this emphasis here, almost in contrast to Section 3.4 where reliable estimation of key secondary and safety effects is brought forward?	Please consider clarifying if it is correctly understood that "reliable" (bias-adjusted/-reduced and accurate) estimates should be provided across ALL efficacy endpoints that are consider relevant for the label. If yes, perhaps consider making this expectation more explicit? Further, clarify if the perceived focus on the primary endpoint in Section 4.1 is in contrast to the statements in Section 3.4.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	218	225	3.42	For adaptive designs where there has been no commonly accepted methods for bias correction in point and interval estimation, it is unclear whether the guidance is recommending to investigate and acknowledge the bias by simulation, or further to develop and apply methods for bias correction.	Suggest to provide clear guidance on the expected actions in situations where specific estimation methods have not been developed. In addition, in line 222, "these should be used" should be modified to discourage naïve point and interval estimates; suggested wording: "these should be used instead of naïve point and interval estimates that don't correct for bias or coverage probabilities".
EFPIA	218	220	3.42	The text "Adaptive design proposals should therefore evaluate bias and variability of treatment effect estimates and provide support of their reliability." should be strengthened to emphasize the importance of quantification of bias and variance in making design decisions.	This phrasing should be strengthened to emphasize the importance of quantification of expected bias and variance or precision in making design decisions.
c4c-S	218	220	3.4	The text is difficult to read.	I recommend to write in bold the sentence: "Adaptative design proposals should therefore evaluate..."
EFPIA	221	222	3.42	Please see suggested change in wording and rationale.	The statement "For some designs, specific estimation methods have been derived with improved reliability, and these should be used." is too broad and should be qualified to state that these methods "should be used when they meaningfully improve the accuracy of estimates of treatment effect."
Ferring Pharmaceuticals	221	222		Could it be clarified whether both median and mean unbiased estimates are acceptable? Many proposed methods for reliable estimation for adaptive designs rely on median unbiasedness.	
MRC Clinical Trials Unit	221	225	3.4	I suggest clarifying the statement to include a cautionary note emphasising that, ideally, the analysis should align with the original design. This alignment is important for maintaining validity and interpretability, and trialists should be aware of its implications. Additionally, it would be helpful to recommend the use of simulation studies as a practical tool to guide decisions when considering deviations from the planned design.	
EFSPI/PSI Regulatory ESIG	224	224	3.4	"bias": in the adaptive design literature there are two concepts of bias: conditional and unconditional. These meaning of them is conceptually very different, and also numerically bias-adjusted estimators can have very different values. Given the focus of the guidance on frequentist inference, it appears indicated that a repeated sampling perspective is preferred. This implies that conditioning on the stage at which a trial stopped does not appear meaningful to quantify bias. Furthermore, many estimators that adjust for conditional bias will shrink the treatment effect dramatically (potentially to the extent that the effect changes direction) if you stop but only barely crossed the efficacy, no matter how extreme the boundary is.	Make specific what type of "bias" is meant. We have a clear preference for unconditional bias, as this allows for a proper frequentist (repeated sampling) interpretation.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFSPI/PSI Regulatory ESIG	233	307	3.5	We appreciate that the guidance dedicates a whole section to trial integrity, and acknowledge that risks for trial integrity and associated mitigation measures within adaptive trial designs are mentioned. However, we feel that one particularly important risk was missed in this section, being the risk of trial personnel trying to predict or guess the subsequent treatment assignments, which can cause several kinds of bias. Especially for open-label trials, which are explicitly mentioned in lines 306-311 to be particularly sensitive to breeches of trial integrity, we feel that the predictability of treatment assignments poses a severe risk to trial integrity which could be addressed in this guideline as well. The permuted block design, which is still the most frequently used randomization method throughout all clinical trials, is known to have a very high proportion of deterministic assignments, thus being very vulnerable to attempts of investigators trying to guess the subsequent treatment assignment (Berger et al. 2021). We believe that a section dedicated to trial integrity within this guideline offers an opportunity to address the shortcomings of the permuted block design especially with respect to open-label trials, and could list other randomization methods as alternatives. The class of maximum tolerated imbalance (MTI) procedures achieve the same degree of control the maximal imbalance of a randomization sequence, while providing a more procedure that is more random, thereby less predictable, hence reducing the risk of selection bias. Examples for these procedures are, e.g. the Big Stick Design (Soares & Wu, 1983), the Biased Coin Design with Imbalance Tolerance (Chen 1999), the Block Urn Design (Zhao & Weng 2011), or the Maximal Procedure (Berger et al. 2003). In addition, we feel that the guideline also should contain an explicit statement to limit access to an open-label trial database, thereby reducing the risk of biased introduced by knowledge of the sponsor. As an example, Higgins et al. (2025) recommend that the "sponsor statistician should be blinded, that is, not have the knowledge of subjects' assignments, until the database is locked and the study is officially unblinded.". In this context, using MTI procedures will also represent an important measure, as this will decrease the risk that the sponsor might be able to predict subsequent treatments based on the current history of treatment assignments within the database. In addition, some recommendation on how to best handle open-label databases, including potential risk mitigation measures, such as restricting parts of the database containing information on the treatment of the patients (or information to deduce the treatment), the use of mock or scrambled information, etc. would be worthwhile additions to this section.	We suggest that the section on "Maintenance of Trial Integrity" should explicitly address the risk of trial personnel predicting future treatment assignments, which can introduce bias, particularly in open-label trials. The guideline could acknowledge the limitations of the widely used permuted block design, which is highly predictable, and recommend alternative randomization methods from the class of Maximum Tolerated Imbalance (MTI) procedures. Additionally, we suggest that the guideline might benefit from including recommendations to restrict access to open-label trial databases to prevent sponsor-related bias, for example by blinding sponsor statisticians until database lock and implementing measures such as limiting access to treatment-related fields or using scrambled data. Combining these measures with less predictable randomization methods would significantly strengthen trial integrity.
EFSPI/PSI Regulatory ESIG	233	307	3.5	[Contined from above] Berger, V., Bour, L., Carter, K. et al. (2021) A roadmap to using randomization in clinical trials. BMC Med Res Methodol 21, 168. Berger, V.W., Ivanova, A. and Deloria Knoll, M. (2003), Minimizing predictability while retaining balance through the use of less restrictive randomization procedures. Statist. Med., 22: 3017-3028. Chen YP (1999). Biased coin design with imbalance tolerance. Communications in Statistics. Stochastic Models 1999; 15(5):953-975. Higgins KM, Levin G, Busch R. (2024Considerations for open-label randomized clinical trials: Design, conduct, and analysis. Clin Trials. 2024 Dec;21(6):681-688. Soares JF, Wu CFJ (1983). Some Restricted randomization rules in sequential designs. Communications in Statistics - Theory and Methods 12:17, 2017-2034. Zhao W, Weng Y (2011). Block urn design - a new randomization algorithm for sequential trials with two or more treatments and balanced or unbalanced allocation. Contemp Clin Trials 32(6):953-61.	
c4c-S	233	307	3.5	Clarify proportionality for smaller or early-phase adaptive studies	The extent of independence and documentation requirements may be proportionate to trial phase and complexity

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
c4c-S	235	284	3.5	The guideline rightly emphasizes confidentiality but lacks practical, quantitative standards for operational bias control and information leakage.	Add operational framework: "Sponsors should quantify risk of information leakage (e.g., probability of inferring interim treatment effects) and demonstrate measures reducing this probability below a pre-specified threshold (e.g., <5%)."
EFPIA	239	307	3.52	While several safeguards to protect trial integrity are described in this section, with special attention given to open-label trials, the guideline misses an important additional measure, being the use of *less predictable randomization methods* .	Suggest having a mention of using randomization methods such as the permuted block design or minimization methods (with suitable primary analysis methods such as re-randomization methods to control type 1 error rate).
EFPIA	239	307	3.52	One other safeguard that was missed is to limit access of the sponsor to unblinding elements of the database of an open-label trial (Higgins et al. 2024)	Suggest including that in an open-label trial, sponsors should limit their access to unblinding elements through establishing a strict access control plan, including segregating duties, using firewalls (such as creating a distinct blinding and unblinding team), and establishing a data handling plan that ensures all data, including Case Report Forms (CRFs), lab reports, and other documents, are captured and stored without treatment assignment information.
c4c-S	239	285	3.5	Access to unblinded interim results can be necessary for adaptation of a trial. Ideally these should only be available to an IDMC, but it may be necessary to share (non-detailed) results to some persons outside this group. However unblinded results should never be available to personnel directly involved in managing and conducting the actual trial, as is stated in the current text. Also the rationale for an adaptation, and thus the underlying data, could be guessed by personnel involved in the trial. This is adequately worded in the current text.	No change
ACRO	249	253	3.5	ACRO welcomes the inclusion of this sentence.	
EFPIA	254	254	3.53	Please see suggested change in wording and rationale.	Please consider rewording the text that currently reads "A fundamental aspect of many types of adaptive designs is the need for some level of access to unblinded interim results." as "A fundamental aspect of many types of adaptive and non-adaptive designs, e.g., when monitored by an IDMC, is the need for some level of access to unblinded interim results." to more accurate convey the general need for access to such efficacy and safety data.
EFPIA	255	255	3.53	It is not realistic that those who have access to unblinded data (e.g., the IDMC) don't have a conflict of interest. For example, IDMC members will always have a conflict of interest, simply because they are paid by the sponsor. It is more important that conflict of interests are disclosed to the relevant stakeholders and that those with a conflict of interest do not have the sole decision-making authority.	Either remove reference to the conflict of interest; write 'limited and managed conflict of interest'; or clearly outline what constitutes a conflict of interest.
EFPIA	255	260	3.53	Even with an IDMC in place, the sponsor is responsible for implementing the adaptation and generally the IDMC's recommendation. As part this, senior sponsor personnel will always have access to some level of unblinded data.	Recognize the involvement of senior sponsor personnel in implementing the IDMC's recommendation and emphasize the importance for having processes in place that prevent the dissemination of unblinded results. Add reference to Section 5.1 of the guidance document that discusses these aspects

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	256	257	3.53	This statement takes a hard stance on sponsor involvement in unblinded data review: "Personnel having access to accumulating unblinded data should generally be independent in the sense that they do not have conflicts of interest or any role in trial activities and are external to the sponsor"	Suggest modifying language to "completely external to the sponsor" or state "independent from the study/asset team?" Blindly accepting DMC recommendations on multi-hundred million dollar commitments ignores financial and operational realities of drug development. There should be opportunity for the sponsor to provide a firewalled plan regarding leadership decision commitees or something of that nature. In addition, it would also be helpful to refer to Lines 602 to 611 that suggest Sponsor's access to unblinded data with appropriate safeguards to maintaining trial integrity.
Regeneron Pharmceuticals, Inc.	257	258	3.5	While Regeneron agrees that IDMCs are vital for maintaining trial integrity for adaptive designs in a Phase 3 setting, we believe that for nonpivotal studies the sponsor should be able to manage adaptive design without an IDMC. To enhance clarity, we recommend that the Council explicitly incorporate language addressing this distinction into the document.	
Takeda	257	258	3.5	An IDMC may or may not be the best candidate for making preplanned interim adaptation decision, depending on the expertise of the committee member. For complex decision making, it will be beneficial to form a dedicated adaptation committee with appropriate expertise. Similar language is available in FDA draft guidance on the use of data monitoring committees in clinical trials (Feb 2024).	To achieve this, an IDMC or a dedicated independent adaptation body should be in place to review unblinded interim data when such access is needed as part of the adaptive design.
EFPIA	260	262	3.53	The phrase "the IDMC can have an additional role of reviewing interim data for the purpose of implementing the planned adaptations", although unclear in meaning, may be read as implying the IDMC should decide whether a planned adaptation should be implemented as an integral part of the prespecified design.	Please consider rewording as "the IDMC can have an additional role of reviewing interim data for the purpose of verifying the continued scientific appropriateness and safety of implementing the planned adaptations" to better characterize the IDMC as a safety check rather than deciding the trial design based on knowledge of interim results.
MRC Clinical Trials Unit	260	263	3.5	IDMC information is unclear and a change from current common practice 1) in row 260 the document says the IDMC should perform analyses rather than the currently common practice of the IDMC receiving a report of analyses done by others. Or does this mean "reviewing interim data analysis results" ? 2) In row 263, there is the suggestion that the IDMC will make the adaptation decision. Currently IDMCs are often advisory.	
IDSWG	261	262	3.5	IDMC should have access to unblinded patient and group-level data	"an additional role of reviewing interim group and patient-level data for the purpose of implementing the planned"
EFPIA	263	265	3.53	Suggest clarifying that, in addition to standard operating procedures and confidentiality agreements, logistical or physical firewalls should also be considered to limit access to unblinded interim results. Implementing these measures would help ensure that only authorized personnel have access to sensitive information and maintain the integrity of the trial.	Propose change to "*Standard operating procedures, ~~and~~ confidentiality agreements, and logistical/physical firewalls should be put in place to limit access to unblinded interim results beyond the IDMC.*"

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	275	278	3.54	Adaptive designs should be chosen to ensure that the design needs are met and operating characteristics optimized through careful quantification of the benefit and risks of the design and its alternatives. Suggesting that adaptive rules should be selected to limit the ability of back calculating the effect size of the trial if an interim decision is known may potentially require the use of a design that is otherwise suboptimal. To maintain balance, the tradeoff associated with taking the approach mentioned should be described. Even with a traditional, frequentist group-sequential trial, knowledge that a trial has continued after a planned interim analysis provides substantial information about the observed interim treatment effect; however, not including an interim analysis results in risks to both participants and trial sponsors that are generally of greater concern.	Please consider revising to state: “However, limiting the information communicated by knowledge of interim decisions may require compromise of other design goals, such as avoiding the enrollment of either smaller or larger populations than needed, or avoiding exposure to treatment arms or doses that appear to be less effective.”
EFPIA	278	281	3.54	We strongly support the statement that "Details of the adaptation rule could be reserved for a specific document rather than the protocol, such as a confidential appendix to the IDMC charter, that is only accessible to designated sponsor personnel separated from the team managing and conducting any aspects of a clinical trial." In Europe in particular, there are currently numerous issues related to the system used for the submission of Clinical Trial Applications (CTIS) for adaptive trials which lead to sponsor personnel being unblinded to the data.	Please add a sentence stating that "Local/regional submission systems may need to be adapted to allow for the use of such separate document so as to ensure the protection of the blind."
EFPIA	281	285	3.54	This text seems to imply misleading sites and particpants about the reasons for sample size changes. Based on later sections the protocol should outline that a sample size adaptation may take place and the maximum increase/decrease. Even if they do not need to know the updated target giving a "made up" reason does not seem totally ethical.	Delete line 283-285 as the sentences prior to this already suggest limiting the potential number of sample size changes so little knowledge could be inferred. It would be helpful to understand if a protocol amendment would be needed in this stiuation where the sample size re-estimation and sample size changes are pre-planned.
EORTC	283	285	3.5	How is it possible not to inform the site of the change of the sample size when this would probably lead to a protocol amendment to be submitted to the local ethics Committees among others?	
EFPIA	286	289	3.55	The requirements and discussions here are unclear and in particular the sufficient size for different stages of the trial. It is unclear if there should be a powered significance test for differences between stages.	Suggest to change scope from "sufficiency of the size ... for assessing … impact of adaptations". Rather point towards: "The sponsor should assess in the design and analyses stage sensitivity analyses to assess and understand heterogeneity between stages."
EFPIA	286	301	3.55	This paragraph contains multiple requirements with a high requirement to document all these discussions and preparatory steps.	It should be clarified for which type of adaptive design, all steps are needed and to which extent this also applies to group sequential designs.
c4c-S	286	291	3.5	The requirement to discuss adaptation implications with regulators is appropriate but lacks timeline guidance.	Specify: “Regulatory engagement on adaptive design proposals should occur before protocol finalization and prior to first patient enrollment, ideally during formal scientific advice or pre-IND/End-of-Phase 2 meetings.”
Takeda	286	288	3.5	Minor suggestions to clarify changes after the adaptation to be evaluated, aligning to the later part of the paragraph (line 293-298).	Sponsors should discuss with regulators at the planning stage the potential implication of the adaptation on trial conduct, including the type of participants to be enrolled after adaptation, and on the interpretation of the results at trial end.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Cancer Research UK Clinical Trials Unit, University of Birmingham	292	293	3.5	The sentence is unclear and could benefit from rewording for clarification.	No change proposed due to uncertainty about the intention of the sentence.
c4c-S	302	305		The guideline notes special risks but lacks actionable mitigation strategies for open-label settings.	Add: “For open-label adaptive trials, sponsors should employ blinded independent outcome assessment and centralized data monitoring to mitigate operational bias.”
c4c-S	302	307	3.5	Especially for an open label adaptive design maintaining trial integrity can be a challenge. The current text can be stricter.	Add a line cf Maria (c4c-S)'s suggestion
ACRO	308	560	4	Challenge: In addition to the examples in the draft guideline, ACRO recommends that the guideline explicitly recognize adaptive designs used in platform, basket, and umbrella trials, as well as biomarker-driven studies that rely on dynamic patient stratification. Inclusion of these examples would align E20 with the evolving design landscape and ensure harmonization with regulatory guidances such as FDA’s 2019 Adaptive Design Guidance and EMA’s 2022 Reflection Paper.	Recommendation: ACRO recommends use of the term 'master protocols' under which to consider platform, basket & umbrella trials, as discussed in the FDA Guidance: Master Protocols: Efficient Clinical Trial Design Strategies to Expedite Development of Oncology Drugs and Biologics Guidance for Industry, Guidance for Industry, March 2022
c4c-S	308	313	3	The section introduces only frequentist methods and defers Bayesian approaches to Section 5.3. This separation risks fragmented guidance for sponsors employing hybrid or Bayesian-frequentist adaptive frameworks.	Add bridging statement: “Although examples in this section are described using frequentist methods, analogous Bayesian approaches that achieve comparable control of error probabilities and interpretability are also acceptable.”
EFPIA	311	313	4.01	The following discussion is general and could be applied to both Bayesian and frequentist methods.	Suggest rewording as “While the discussion focuses on designs using frequentist approaches for statistical analysis, many of the considerations apply to both Bayesian and frequentist methods.”
EUCROF - EU CRO	314	370	4.1	The primary endpoint of a clinical trial could be a safety endpoint and early trial stopping could be based on safty considerations that had been predefined. We think an explanation why this section is based on efficacy endpoints only, is missing, or, alternatively, safty endpoints should be addressed as well.	
c4c-S	314	370	4.1	Could better distinguish between statistical futility and safety-based termination	Add clarification distinguishing “futility for lack of efficacy” versus “early stopping for safety or external evidence.”
EFPIA	315	316	4.11	This is a very general statement, as any trial can be stopped based on accumulated data, especially for safety reasons or ethical reasons, such as preventing subject allocation to an apparently inferior control. It is too ambiguous to use this as an introduction to a planned decision-making paradigm of a pre-specified sequence of analyses designed to optimize decision making.	Remove: "During the conduct of a clinical trial, accruing data can provide information that makes it no longer appropriate to continue the trial. To address this, "
EFPIA	318	319	4.11	An adaptive clinical trials can also be stopped for safety reasons.	(…) anticipated rules for stopping when there is compelling evidence of efficacy (stopping for efficacy), ~~*or*~~ when the trial is unlikely to demonstrate efficacy (stopping for futility)*, or for safety reasons* .

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	319	323	4.11	"A clinical trial design that allows such sequential analyses for early efficacy stopping based on accumulating observations of groups of participants at pre-specified points throughout the trial is called a group sequential design."	We suggest modifying this to: "A clinical trial design that allows such sequential analyses for early stopping for efficacy or futility based on accumulating observations of groups of participants at pre-specified points throughout the trial is called a group sequential design."
c4c-S	322	322	4.1	I think to highlight a term is needed to clarify the text.	I would write in bold "group sequential design".
c4c-S	323	359	4.1	Difficulties in understanding the subsection.	I would create a sub-subsection named 4.1.1. Stopping for efficacy
EFPIA	331	334	4.12	The text that reads "In addition, methods for calculating the primary treatment effect estimate and associated confidence interval that adjust for the interim analyses should be planned to limit bias and improve performance on measures such as the mean squared error (Section 3.4)" should directly acknowledge and address the bias-variance tradeoff. The focus should be on maximizing the probability that the trial result is correct overall, a combination of minimizing false-positive and false-negative results. In many contexts, the risk of a false-negative result is a substantial component of the risk of "getting the wrong answer," and the goal of minimizing this risk may appropriately motivate the choice of estimation procedures that result in some bias. The key is that the tradeoff between different measures of validity in estimation is transparent, quantified, and justified in terms of maximizing the trial's potential to support improvements in treatment and health of the affected population.	Please modify to focus on the goal of maximizing the probability that the trial result is correct overall, a combination of minimizing false-positive and false-negative results.
EFPIA	335	338	4.13	The text, stating "A trial that is stopped early for efficacy will provide less information (e.g., because of a smaller sample size and/or shorter duration of follow-up) for the evaluation of safety, important secondary efficacy endpoints, and relevant patient subgroups, which are important for the overall benefit-risk assessment" has important underlying assumptions and, if these are not communicated clearly, risks being misinterpreted in an overly broad manner. For example, a smaller adaptive trial that allocates a larger fraction of the total population to the arm or dose that ultimately is of the greatest interest may, in fact, provide more information of importance than a larger non-adaptive trial that allocates participants equally across arms.	Please consider rewording and expanding as "For any particular trial design, adaptive or non-adaptive, a trial that is stopped early for efficacy will generally provide less information than one that proceeds to its maximum N." The next existing sentence then follows naturally, "Therefore, the timing of interim analyses should be selected such that the sample size is large enough and the duration of follow-up is long enough to ensure sufficient information is available for decision-making" and no modification of it is needed.
c4c-S	335	335	4.1	To change the word "will"	I think "would" it's more precise in this sentence.
c4c-S	335	344	3	The guideline rightly notes limited safety information but gives no operational mitigation measures.	Add: "When early efficacy stopping is possible, sponsors should pre-specify minimal cumulative exposure thresholds and post-trial follow-up plans to ensure adequate safety data."
c4c-S	335	342	4.1	The guideline rightly notes limited safety information but gives no operational mitigation measures.	Add a line cf Maria (c4c-S)'s suggestion
c4c-S	337	347		Consider highlighting that in vulnerable populations such as children, early stopping for efficacy or futility requires additional ethical scrutiny. Continuing a control arm after strong efficacy signals can be ethically challenging. Conversely, premature stopping may jeopardise understanding of long-term safety or developmental outcomes.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
ACRO	338	338	4.1		Specifically, we recommend the addition of a new sentence in line 313: “Adaptations may include adaptive designs implemented under master protocols (e.g., platform, basket, umbrella trials) and biomarker-driven studies employing dynamic stratification.”
ACRO	338	338	4.1	ACRO supports the emphasis on ensuring adequate sample size when considering the timing of interim analysis in order to support safety analyses.	
EFPIA	340	340	4.13	Suggest improving wording for better clarity.	Change "decision-making" to "benefit-risk assessment".
EFPIA	345	349	4.13	This sentence suggests minimal benefits of GSD in other cases, for example, earlier drug availability to patients in unmet medical need situations.	Change "(e.g., the primary endpoint is survival)" to "(e.g., the primary endpoint is survival or there is an unmet medical need for patients)".
EFPIA	345	349	4.13	There are multiple important motivations for using adaptive designs with early stopping for efficacy, e.g., there is no effective therapy for a condition that is not life-threatening but causes substantial morbidity or suffering. To avoid an overly narrow interpretation of the point being made, please consider revising this sentence as suggested.	Please consider revising this sentence to read “Interim analyses with the potential for early stopping are often considered in circumstances where there are compelling ethical reasons (e.g., the primary endpoint is survival), and efficacy stopping rules typically require highly persuasive results in terms of both the magnitude of the estimated treatment effect and the strength of evidence of an effect.” Deleting “Furthermore” and “more” makes this a statement about one utilization of these approaches rather than suggesting the use is limited to this setting.
EFPIA	350	359	4.14	The text is a bit ambigous as to in case a trial is stopped based on interim results, then whether it is the results of the interim analysis that are the trial results that should be evaluated (including secondary endpoints and safety data) or if it is the analysis performed on the data including the patients that were ongoing in the trial at the time of the interim and continue their data contribution afterwards. Clarity on this point would be of great value	
EFPIA	350	359	4.14	In the setting of “overrun” with the arrival of additional outcome data after an interim decision to stop a trial for efficacy, while we agree wholeheartedly with the recommendation that all data be completely and transparently reported, it is important that the prespecified design define which dataset—the dataset that resulted in the decision to stop or the complete dataset—is to be considered the primary trial result. Either choice is defensible, however: (i) if the dataset that motivated the stopping decision is considered primary, then there should be expected to be some regression to the mean in the final data and a “loss” of nominal statistical significance should not alter the overall conclusion; and (ii) if the complete dataset is considered primary then there is a non-zero but usually small chance that the final result may not meet the original stopping threshold and the trial must be interpreted as negative. In either case, realistic simulations–implementing various design decisions—can be used to understand the magnitude and likelihood of these occurrences and the effects on error rates, and support the ultimate design choice.	Specifically, we suggest that the following be inserted in place of the sentence currently on lines 356-359: “When such “overrun” is possible or occurs, it is critically important that all data be completely and transparently reported. Moreover, during the design of the trial, it is important that the prespecified design defines which dataset—the dataset that resulted in the decision to stop or the complete dataset—is to be considered the primary trial result. Either choice is defensible, however: (i) if the dataset that motivated the stopping decision is considered primary, then there should be expected to be some regression to the mean in the final data and a “loss” of nominal statistical significance should not alter the overall conclusion; and (ii) if the complete dataset is considered primary then there is a non-zero chance that the final result may not meet the original stopping threshold and the trial must be interpreted as negative. In either case, realistic simulations–implementing various design decisions—can be used to understand the magnitude and likelihood of these occurrences and the effects on error rates, and support the ultimate design choice.”

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Cancer Research UK Clinical Trials Unit, University of Birmingham	360	362	4.1	The suggestion is that futility rules should be nonbinding, but the same has not been said for efficacy. Should it be interpreted from the guidance that efficacy stopping rules should be binding?	
c4c-S	360	370	4.1	Difficulties in understanding the subsecion.	I would create a sub-subsection named 4.1.2. Stopping for futility
IDSWG	364	375	4.1	Detailed inclusion of futility stopping rules (timing and boundary) in the protocol could very well hinder trial integrity if the trial continues after the futility interim (eg., changes to the enrolled population (shifts), period effects impacting protocol adherence, retention, and outcome assessments, etc.).	Perhaps futility stopping rules (timing and boundary) are details that could be prespecified in a less public document (eg., the interim statistical analysis plan) in order to maintain trial integrity. Without explicitly allowing for this, member states will continue to ask for inclusion in the protocol.
Teva Pharmaceuticals	366	368	5.1	"This means that the futility stopping criteria serve as guidelines that can be deviated from based on the interim results without increasing the Type I error probability" • Further clarity would be helpful. Specifically, it is not clear if the use of binding futility rules is not recommended or if it should be considered under certain conditions.	
EFSPI/PSI Regulatory ESIG	371	427	4.2	The draft guidance seems to provide more room for an adaptive design than a GS one, even in the case the adaptive nature in 'only' on the sample size. This is also emphasized in Section 4.1, lines 345 to 349, where the bar for the inclusion of early stopping rules is high. Another example of difference in tone between GSD and adaptive design with SSR occurs later, lines 529 to 531. "[...], RAR designs are susceptible to bias and inflation of the Type I error probability in the presence of overall time trends." While the statement is understood, this concern seems to not only be appropriate for RAR, but also for adaptive designs with SSR, although arguably the problem is less pronounced. However, the adversity towards time trends might be seen as a motivation to 'start high' in a GSD rather than an adaptive design, all other things being unimpactful.	Consider clarifying if the perceived unequal treatment in the draft guidance between GSD (setting the bar for early stopping for efficacy very high) and adaptive designs (with SSR only) is intentional.
c4c-S	371	386	3	Clear rationale given, but the text underplays regulatory expectation for documentation of simulation assumptions.	Add requirement: “A detailed simulation plan should accompany submissions, including assumed control response rates, variability, and sensitivity analyses demonstrating robustness.”
Regeneron Pharmceuticals, Inc.	371	427	4.2	Interim analysis is generally a well understood and well accepted adaptive design. Regeneron recommends the guidance discuss the trade off between a pure sample size adaptation design and a conventional group sequential design. To enhance clarity, we recommend the Council incorporate language to outline the situations when one should opt for sample size re-estimation instead of group sequential.	
EFPIA	374	375	4.21	Detailed inclusion of futility stopping rules (timing and boundary) in the protocol could very well hinder trial integrity if the trial continues after the futility interim (eg., changes to the enrolled population (shifts), period effects impacting protocol adherence, retention, and outcome assessments, etc.).	Perhaps futility stopping rules (timing and boundary) are details that could be prespecified in a less public document (eg., the interim statistical analysis plan) in order to maintain trial integrity. Without explicitly allowing for this, member states will continue to ask for inclusion in the protocol.
c4c-S	374	385	4.2	The text is difficult to read.	I reccommend to add numbers in the different sources of uncertainties.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Invents consortium - EU Horizon project	374	378	4.2	The description of "nuisance parameter" in lines 374–378 does not reflect the broader set of unknown parameters that may have an influence on the type 1 error rate of adaptive designs and therefore need to be considered in simulation studies (lines 656–679). Beyond the usual examples (SD for continuous outcomes and control response rate for binary outcomes) the operating characteristics of adaptive designs are affected by additional unknown parameters.	For simulation studies for adaptive designs, the term "nuisance parameters" should mean all unknown features of the data-generating model not fixed by the null hypothesis and trial design. Examples are: (i) multi-arm trials - when testing one arm, the outcome distributions of other arms (e.g. characterised by their effect sizes and variances) are nuisance parameters and can influence the Type I error rate (both, the familywise error rate and the per-comparison error rate); (ii) adaptive enrichment trials: when testing one subgroup, effect sizes in the other subgroups are nuisance parameters; (iii) designs with adaptations based ob information from secondary or early outcomes: when testing the primary endpoint, effects in the secondary/early endpoints are nuisance parameters; (iv) time-to-event settings: accrual distributions have an impact on the censoring distribution and therefore must be treated as nuisance parameters.
EFPIA	375	380	4.21	A non-technical discussion about "nuisance parameters" is a bit dangerous in my view. The responder rate in the control arm may be a "nuisance parameter", but obviously monitoring it to adjust sample size of the trial say, will lead to serious bias, because it is a crucial element of the treatment effect estimate. Thus, casual readers of the guideline might interpret this as a license to monitor the control responder rate for early stopping. I cannot think of any situation where this is a feasible strategy, because unblinded sample size rveiwe will always be better when proper adjustement is done.	Rewrite this or delete the example. If rewrite, one could say: " However, it is crucial to formally define and differentiate nuisance parameters from those related to treatment effects. For example, while overall responder rates derived from blinded data do not reveal treatment assignment, they may still provide indirect information about treatment effects, such as risk differences."
Teva Pharmaceuticals	376	433	5.2	• Please consider providing clarification on the models. For example, consider whether the models developed for trial simulation to support the justification of adaptation choices under ICH E20 need to be fully qualified (as discussed in ICH M15). Alternatively, please clarify whetheher fit-for-purpose models can be used as evidence to justify adaptive design features, provided their assumptions, verification, and sensitivity analyses are transparently described. • Please consider that controlling Type-1 error, sample size and timing of IA can remain the same issue in non-adaptive trial studies. Accordingly, it would be helpful to clarify whether the sponsor should demonstrate superiority of propose of the adaptive design over other options e.g. by simulation.	
IDSWG	381	385	4.2	This should appear at the beginning of Section 4.2 " Another source of uncertainty at the planning stage are assumptions about the anticipated treatment effect size. In cases where there is justification based on residual uncertainty (e.g., after appropriate exploratory trials; see Section 3.1), sponsors may consider a sample sizeadaptation based on an interim treatment effect estimate. The goal would be to ensure sufficient power under a range of plausible and clinically meaningful treatment effect sizes.	
EFPIA	385	385	4.21	Usually it is not meaningful to power for plausible (versus clinically meaningful) treatment effect sizes.	Remove "plausible".
EFPIA	391	394	4.22	"Adaptations to the sample size based on nuisance parameter estimates should be carried out using blinded data as this approach does not incorporate information about treatment assignment, thus minimizing risks for trial integrity." However, per lines 375-378: "Examples of nuisance parameters include the standard deviation of a continuous outcome and the probability of response of the control arm for a binary outcome…" In case of binary outcome, the sample size adaptation can be done using unblinded data.	Therefore, we suggest to change Lines 391-394 to: "Adaptations to the sample size based on nuisance parameter estimates should be carried out using blinded data if possible as this approach does not incorporate information about treatment assignment, thus minimizing risks for trial integrity."

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Cancer Research UK Clinical Trials Unit, University of Birmingham	392	394	4.2	The recommendation is to use blinded data to adapt the sample size for nuisance parameters. If the nuisance parameter is a difference between predicted and observed control outcomes then how can sample size adaptation be performed without being aware of the control rate if blinded. Could the guidance clarify suitable approaches. This is confusing given the statement in line 404 that states adaptations are based on unblinded data?	
EFPIA	392	392	4.23	In line 377 control response rate is defined as a nuisance parameter but in line 392 it is suggested that it is preferable to investigate nuisance parameters in a blinded fashion. It's unclear how this would be possible.	Revise line 377 or provide additional advice on how it is possible to assess control response in a blinded fashion.
EFPIA	392	392	4.23	The word 'should' here is excessively restrictive, as would be the case when estimating a nuisance parameter in multi-arm studies. It also seems to be contradictory to previous paragraphs which use response rate at control group (a nuisance parameter) for adaptation.	Replace "should" with "may".
EFPIA	392	394	4.23	(See previous comment for lines 375-380) Quantities derived from blinded data may tell you nothing about treatment assignment, but they may tell something about treatment effects. For example, the overall responder rate call tell something about risk difference. Without a proper formal definition of "nuisance parameter", it remains unclear what this staement is supposed to say.	Adaptations to the sample size in clinical trials should either be based on estimates derived from blinded data to avoid incorporating information about treatment assignment, thereby minimizing risks to trial integrity, or should properly be adjusted when based on unblinded data.
EFPIA	392	394	4.23	The current text misses the opportunity to discuss the risks of the approach of using nuisance parameter estimates from data aggregated across treatment groups (“blinded” data). While this may minimize the risk for trial integrity—an advantage that should be explicitly stated—it may also substantially increase the risk of making an erroneous interim decision as the estimates for the nuisance parameters based on pooled data are influenced by the treatment effect that is not unaccounted for. For example, the sample size may be increased unnecessarily when there is a larger treatment effect because the pooled estimate of the variance is inflated due to the treatment effect.	Please consider replacing the word “should” on line 392 with “may” and inserting the following text on line 394, between the two existing sentences: “While this approach may facilitate protecting the integrity of the trial, it risks introducing bias in the sample size reestimation as the estimates for the nuisance parameters based on pooled data are influenced by the treatment effect that is not accounted for. For example, the sample size may be increased unnecessarily when there is a larger treatment effect because the pooled estimate of the variance is inflated due to the treatment effect.”
EFSPI/PSI Regulatory ESIG	392	394	4.2	"Adaptations to the sample size based on nuisance parameter estimates should be carried out using blinded data as this approach does not incorporate information about treatment assignment, thus minimizing risks for trial integrity"	Perhaps consider adding some clarification if this is always generally feasible, eg, if there is high uncertainty about the response in the control group.
EFPIA	398	403	4.23	Please see suggestion in next column on this paragraph	We recommend to add additional context: “In some cases, conventional analysis methods that would apply in non-adaptive designs can be used for the primary analysis if there is justification (e.g., in a reasonably sized two-arm superiority trial with a continuous endpoint, where adaptations such as blinded sample size re-estimation or early stopping with alpha-spending functions preserve the null distribution as well Type I error control). In other cases (e.g., a two-arm non-inferiority trial with a continuous endpoint, where adaptations such as sample size re-estimation may alter the null distribution due to the fixed non-inferiority margin), the use of these conventional methods may lead to an increase in the Type I error probability and different approaches are needed.”

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
IDSWG	398	401	4.2	The sentence is not clear "In some cases, conventional analysis methods that would apply in non-adaptive designs can be used for the primary analysis if there is justification (e.g., in a reasonably sized two-arm superiority trial with a continuous endpoint)."	
IDSWG	398	403	4.2	These sentences are not clear to me. What conventional methods are we talking about? In some cases, conventional analysis methods that would apply in non-adaptive designs can be used for the primary analysis if there is justification (e.g., in a reasonably sized two-arm superiority trial with a continuousendpoint). In other cases (e.g., a two-arm non-inferiority trial with a continuous endpoint), the use of these conventional methods may lead to an increase in the Type I error probability and different approaches are needed.	
EFPIA	400	403	4.23	The guideline provides examples of continuous endpoints for both cases, what about the binary endpoint case or survival endpoint case? Do we have the same considerations for superiority and non-inferiority? This is important because many clinical trials have these types of endpoints.	Please add a reference and consider adding examples that use different types of endpoints such as binary.
EFPIA	401	401	4.23	The guidance only mentions two-arm non-inferiority trial with continuous endpoints as an example, whereas this challenge of type I error rate inflation is also equally prevalent in equivalence trials also handling binomial endpoints.	Suggest using the phrase in the example "(eg., a two-arm non-inferiority or equivalence trial),"
Takeda	404	406	4.2	An IDMC may or may not be the best candidate for making preplanned interim adaptation decision, depending on the expertise of the committee member. For complex decision making, it will be beneficial to form a dedicated adaptation committee with appropriate expertise. Similar language is available in FDA draft guidance on the use of data monitoring committees in clinical trials (Feb 2024).	Trials with sample size adaptation based on interim effect estimates should use an IDMC or other adaptation body and adequate processes to maintain trial integrity, given that the adaptations are based on unblinded data.
EFPIA	413	418	4.25	Here and in other sections (see, e.g., lines 152-154, 363-366), the guidance states that strict adherence to the anticipated adaptation rule is not required, provided that statistical integrity, such as Type I error rate control, is maintained. We believe the rationale behind this stems from the desire to avoid interrupting recruitment while a decision is considered and to allow deviations from the anticipated timing of the interim analyses (lines 328-329), and, more importantly, from the need to collect sufficient information to support the overall benefit-risk assessment. This last point leaves a door open for substantial divergences between pre-specified and realised adaptations, as alluded to by the use of the word "anticipated" in the adaptation rule and by the text in section 6.2 that requires describing the adaptation as it actually happened (as opposed to as planned), along with a rationale for this. Consequently, whilst an adaptive design can be compellingly justified in its planning stage, its delivery could be very far from that justification. To prevent the advantages and 'raison d'être' of adaptive designs from being easily undermined during the course of the trial, we suggest that the preference for non-binding rules be more explicitly caveated. This clarification could be included as part of the general principles in the document.	Suggest editing the sentence in line 163 (section 3.2) to read "and outline factors that may lead to such deviations as well as the consequences of deviating on the choice of adaptive elements included". It is also suggested that the rationale for preferring non-binding rules that appears in principle across the document be discussed more clearly as it is somehow missing why or when this would be acceptable or desirable. Additionally, we suggest that the potential impact of deviations from the anticipated adaptation rule should be evaluated through simulations, similar to the sensitivity analyses performed for non-binding futility rules to assess their effect on the type I error rate and power. For this we suggest that line 710 (bullet point 8) is modified as follow: "This should include a detailed discussion of the proposed adaptive design and its estimated operating characteristics under various scenarios. Such scenarios should include the impact of possible deviations from the planned rules."
c4c-S	413	427	4	The draft advises bias adjustment but omits quantitative acceptability thresholds.	Add: "Simulation-based evaluation should quantify expected bias and coverage; bias exceeding 5 % of true effect should prompt justification or design modification."
Takeda	415	418	4.2	An IDMC may or may not be the best candidate for making preplanned interim adaptation decision, depending on the expertise of the committee member. For complex decision making, it will be beneficial to form a dedicated adaptation committee with appropriate expertise. Similar language is available in FDA draft guidance on the use of data monitoring committees in clinical trials (Feb 2024).	Still, the anticipated adaptation rule should be pre-specified to facilitate the evaluation of trial operating characteristics (e.g., expected sample size and power) and ensure that the IDMC (or other adaptation body) understands and is in agreement with the anticipated adaptation rule.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
BSWG	420	423	4.3	Population Selection does not discuss additional safety concerns. "For example, a treatment may be expected to benefit a certain targeted subset of the overall population without additional safety concerns…"	
EFPIA	425	426	4.26	suggest providing additional explanation or description of the some cases	"…, a decision would be made at the end of the trial to evaluate the treatment effect only in the overall population or in both the overall population and the targeted subpopulation."
EFSPI/PSI Regulatory ESIG	426	426	4.2	"In some cases, methods are available that adjust estimates to reduce or remove bias associated with the adaptation and these are preferred."	This formulation begs the question "And what if no methods are available?" What does "reliable adjustment" mean in such a case? We suggest to make that precise.
IDSWG	426	427	4.2	Can examples be provided of what methods are being pointed at here? In some cases, methods are available that adjust estimates to reduce or remove bias associated with the adaptation and these are preferred.	
EFPIA	428	428	4.3	Within Section 5.3, most of the texts address methods of borrowing information. As an additional example, one could consider the use of a non-informative prior, such as for safety monitoring or for specifying a stopping rule for a treatment arm.	Please consider adding additional example as suggested.
EORTC	428	428	4.3	For many treatments , the optimal population is not clearly identified at the time of licensing i.e. immunotherapy for cancer. Post licensing research is critical for optimisation and effectiveness. Therefore, the scope (section 1) should be further expanded in that respect to post licensing trials	
EFPIA	429	429	4.31	Population selection is a relevant approach to addressing uncertainty on the optimal patient population. However, it is unclear from this section, how integrity could be maintained in a trial with population selection. Investigators and sponsor will typically be aware of the enrichment decision and can thereby infer some considerations on treatment effects. Guidelines on how to control integrity for this type of adaptation should be provided.	Add considerations to maintaining trial integrity for population enrichment trials.
EFPIA	429	513	4.31	Adapting the population or the treatment conditions in the study consitutes a selection of the clincial question of interest and the estimand. The study documents should have no ambiguity on the estimand.	The guidance should clearly state the need that any adaptations to or selections of estimands have to be documents accordingly in the study protocol.
c4c-S	429	473		Adaptive population selection should explicitly consider paediatric (sub)cohorts defined by age, maturation stage, or developmental pharmacology. Interim adaptations could allow refinement of age ranges or weight bands once early data clarify dose-response or safety differences. The guideline could encourage simulation scenarios reflecting these hierarchical populations.	
EFPIA	457	457	4.32	The enrichment discussion primarily involves experiments with two population subgroups (such as biomarker positive and negative). This section should also include a discussion of basket trials, where multiple indications or subgroups might be considered, e.g., in rare oncologic disease settings. Within this setting, the use of Bayesian models or similar frequentist strategies should be discussed, particularly as the usual discussion of bias becomes problematic. For example, it is well known that, even when raw estimates are unbiased in isolation, the highest raw estimate from a group of raw estimates is biased upward, so the use of individually unbiased estimators does not guarantee unbiased estimates after selection of the population of interest. Hierarchical models are intended to address this form of bias and produce superior estimates.	Please include a discussion of basket trials, the considerations around obtaining treatment estimates when multiple populations are being studied, and the value of hierarchical modeling in this setting.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
c4c-S	458	486	4.3	The text is difficult to read.	I would write a list of the aspects that should be included in the planning of the population selection of adaptative designs.
Invents consortium - EU Horizon project	465		4.3		Exactly what is meant by a Type I error in the population selection setting should be defined
EFPIA	472	473	4.33	Please see comment above re lines 331-334.	Please consider rewording to read: "…and estimates that reduce mean-square error or bias should be considered if the evaluation of the conventional treatment effect estimates demonstrates that the likely magnitude of the bias is sufficient to risk compromising the interpretation of the trial."
IDSWG	472	473	4.3	What methods need to be used to obtain adjusted estimates? The reliability of the treatment effect estimates in the different populations should be evaluated, and adjusted estimates that reduce or remove bias should be considered.	
c4c-S	474	483	4	The guideline calls for rationale but lacks explicit criteria for acceptable biomarker or threshold justification.	Insert: "For biomarker-defined subpopulations, sponsors should provide data-driven threshold justification and sensitivity analyses assessing robustness to threshold shifts of ±10 %."
c4c-S	478	492	4	The section appropriately introduces adaptive treatment selection but limits examples to two doses or two treatments, potentially underrepresenting multi-arm or platform designs.	Add statement: "The principles described also extend to multi-arm or platform confirmatory trials where multiple experimental treatments or doses are evaluated under a shared control and adaptive selection is used to continue only the most promising arm(s)."
EFPIA	480	483	4.34	In the setting of an adaptive "enrichment" trial with prespecified population selection criteria, the requirement that the adaptive trial design must ensure "…that the trial will provide adequate information on the benefit-risk profile in the complementary subpopulation" is overly burdensome and far beyond what is required of a traditional non-adaptive trial. If a positive trial, after the selection of a smaller target subpopulation, is intended only to support use of the therapy in that subpopulation, determining the benefit-risk profile in a different population is unnecessary. A sponsor that runs a non-adaptive confirmatory trial to support regulatory approval in one population does not have to (and rarely does) identify the risk-benefit of the treatment in a complimentary population for which the treatment is not intended.	Please revise to avoid implying the requirement for detailed benefit-risk data on a sub-population that has been eliminated from consideration using an enrichment trial design.
ACRO	484	486	4.3	Challenge: This section would benefit from an example about how to precisely define the ranges/thresholds of "continuous" baseline characteristics of subpopulations.	Recommendation: At the end of line 486, we suggest adding: "Real-world data can be explored, in the planning phase, to determine the common ranges of such continuous/non-binary characteristics (e.g. age ranges, lab values, etc.) in subpopulations of interest, which could help refine respective eligibility criteria in order to optimize benefit-risk profile, and estimate feasible and indicative subpopulation sample sizes for such adaptive designs."

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
IQVIA	486	486	4.3	Providing an example of how ranges for “continuous” baseline characteristics of subpopulations could be defined would address a common hesitation we see among protocol developers. While it is up to sponsors and collaborators involved to determine their methods, providing this text would advance the industry’s confidence with applying real-world data sets for better study designs.	Add at the end of the this current sentence: "If the baseline characteristic that may be used to define subpopulations is not binary in nature, justification should be provided at the planning stage for any threshold(s) used to define the subpopulations."Real-world data can be explored in the planning phase to determine the common ranges of non-binary characteristics (e.g.: age ranges, lab values, COA scores, etc.) of subpopulations of interest; which could help refine/fine-tune respective eligibility criteria, for optimizing benefit-risk profile, and estimate feasible and indicative subpopulation sample sizes for such adaptive designs”
c4c-S	487	498	4.4	The section appropriately introduces adaptive treatment selection but limits examples to two doses or two treatments, potentially underrepresenting multi-arm or platform designs.	Add a line cf Maria (c4c-S)'s suggestion
Breakthrough T1D	488	489	4.4	For conditions with no cures, such as T1D, paired combinations of therapies with distinct and complementary mechanisms of action may be necessary, for example utilizing disease modifying immunotherapies alongside cell replacement therapies. In these scenarios, dosing regimens may vary widely, including sequential or concurrent administration, and require adaptation as evidence accumulates during the trial. The EMA should request the ICH to consider including a dedicated subsection in the guidance to address this emerging trial paradigm, which is expected to become increasingly relevant in diseases with high unmet medical need.	
EFPIA	488	488	4.41	Treatment selection may also impact randomisation ratio to placebo i.e. in situations where only one active arm is dropped, one will need to decide whether a 1:1 allocation vs. placebo is maintained or whether there will be a 2:1 allocation going forward. Considerations on this should be included into this section, as well as considerations on how to maintain integrity for trials with treatment selection at interim analysis.	Add considerations to change of allocation ratio in the section on treatment selection.
EFPIA	493	493	4.41	Suggest improving wording for better clarity.	Replace "conceivable" with "beneficial".
c4c-S	499	513	4.4	The text is difficult to read.	I recommend to add numbers to the aspects included in the adaptative treatment selection, i.e., 1-specification of the treatments, 2-the decisions to be made, etc.
c4c-S	502	503	5	The requirement to “manage participants” is general and operationally vague. It should specify safety follow-up and ongoing data use.	Replace with: “Sponsors should prespecify how ongoing participants on a discontinued treatment will be managed, including safety follow-up duration, continuation criteria, and inclusion/exclusion of their data in the final analysis.”
c4c-S	502	503	4.4	The requirement to “manage participants” is general and operationally vague. It should specify safety follow-up and ongoing data use.	Add a line cf Maria (c4c-S)'s suggestion
c4c-S	504	505	5	The section correctly highlights the need for Type I error control but lacks reference to accepted statistical frameworks (e.g., closed testing, combination functions, multiple-stage p-value combination).	Add: “Acceptable methods include combination-function, conditional-error, or closed-testing frameworks ensuring strong Type I error control across adaptive stages.”

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	505	507	4.42	The adverse consequences on trial efficiency associated with allowing flexibility in adherence to adaptation rules need to be enumerated and discussed; the current text is silent on these issues. Since allowing flexibility while maintaining desired operating characteristics, e.g., type I error control, generally requires more stringent thresholds for declaring efficacy than if the prespecified adaptation rules can be assumed to be followed, the flexibility can result in a requirement for a larger sample size or reduced power. A design that requires adherence to prespecified adaptation rules also means that, if a decision is made to deviate from those rules, e.g., in response to data patterns in secondary or safety endpoints, then the designed operating characteristics may not be preserved. Thus, the decision to allow—or not allow—such flexibility within the prespecified design should be motivated by an explicit consideration of the advantages and disadvantages of both approaches, rather than assuming allowing flexibility is uniformly preferable. Please see related comment re lines 150-154 in Section 3.2.	Please see suggestion related to lines 150-154 in Section 3.2, namely: The text should be modified to clarify the two competing considerations, namely (1) the ability to maintain valid inference despite flexibility in adherence to prespecified decision rules; and (2) the ability to optimize operating characteristics conditional on the assumption the prespecified rule will be followed. Since both goals cannot, in general, be realized simultaneously, the choice regarding the analysis method should be based on quantitative or semiquantitative assessment of the relative risks of the two approaches, rather than one approach being "generally" preferred.
Invents consortium - EU Horizon project	505		4.4		Exactly what is meant by a Type I error in the treatment selection setting should be defined
c4c-S	506	509	5	The text recommends flexibility but should clarify boundaries between acceptable flexibility and ad hoc decision-making.	Add: "Flexibility should be implemented within pre-defined adaptation corridors or decision ranges documented in the protocol or SAP to prevent post hoc operational bias."
c4c-S	506	509	4.4	The text recommends flexibility but should clarify boundaries between acceptable flexibility and ad hoc decision-making.	Add a line cf Maria (c4c-S)'s suggestion
c4c-S	511	513	5	The draft notes bias but omits minimum reporting expectations for adjusted estimates.	Add: "Sponsors should present both naïve and bias-adjusted estimates (e.g., via shrinkage estimators, conditional estimators, or bootstrap-based corrections) with justification for chosen adjustment methods."
EFPIA	514	560	4.5	In discussing response adaptive randomisation, it would be helpful, if the guidance could clarify the situations, where those approaches are considered of relevance, which would be given by very rare indications, paired with strong expected efficacy benefits. Trials in this situation should treat and inform and hence the design aims to establish statistically significant superiority while maximizing number of treatment successes within the trial. While unbiased estimation is always of interest, it may not be the primary aim of the trials. Challenges related to power and potential time-trends remain but would need to be assessed vs. the concern of conducting the trials in the small populations. Also note that response-adaptive randomisation (RAR) trials could be a primary application of Bayesian Inference. They could be of particular relevance in hard to recruit pediatric situations, where information from adults may be extrapolated through a Bayesian model. It may be worthwhile to rather refer here to Bayesian Inference instead of pointing to possibly inappropriate combination tests, which would lose efficiency due to suboptimal planning of stagewise weighting.	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	515	560	4.5	Guidance has a distinct subsection on Response Adaptive Randomization (RAR) • Briefly mentions Covariate Adaptive Randomization, but then does not address the topic • Mentions that RAR designs are susceptible to inflation in Type I error probability in presence of overall time trends. This is not always the case as the control of the Type 1 error rate depends on many factors, such as the background response model used during the adaptation process, the size of the burin, and the balanced allocation used during the burnin stage. Using an efficient restricted randomization method at the burin stage can give a well-controlled type I error rate even while using RAR. • Time trend suspectibility mentioned (which however is no problem exclusive to RAR) • Indirectly recommends few to one timepoints of adaptations, but gives no (good) reason why • States that RAR can lead to patients being allocated to arms with a bad benefit-risk profile → obvious solution: use combined efficacy safety-endpoint for RAR? • Encourages the use of weighted test statistics as analysis method, claiming that randomization tests are generally less powerful (without any literature evidence) • The document does not address the topic of Covariate-Adjusted Response Adaptive (CARA) designs which is distinct from RAR, but can be more efficient due to covariate adjustment. Trials such as I-SPY2 has used a Bayesian CARA method and can be used as a real life example trial that has been successful. • The guidance mentions that RAR could lead to insufficient sample size to support decision make on a treatment. For RAR the overall sample size is fixed and pre-trial simulation is performed to check for the maintenance of the overall power and the type I error rate control. Therefore this comment is suggested to be removed. • Encourages the use of weighted test statistics as analysis method, claiming that randomization tests are generally less powerful (without any literature evidence) • For some reason mentions the deterministic play-the-winner rule at the end of the section (however without mentioning the trial explicitly, and also stating that this is not randomization)	Please modify as suggested as missing in each sub-points
EFSPI/PSI Regulatory ESIG	515	516	4.5	It is stated here that in RCTs, "participants are typically allocated to treatment arms according to fixed randomization probabilities". The term "fixed randomization probabilities" is somewhat ambiguous here. In our opinion, it would be important to distinguish here between 1.targeted allocation probabilities (which are constant for most trials, e.g. 0.5 for each arm in a 1:1 randomized trial) 2.conditional allocation probabilities, i.e. the probability for a given patient to receive a given treatment conditional on the previous treatment assignments – these are not constant by design for any restricted randomization procedure, but are changed in order to meet some balance prerequisite (i.e. are set to 0 or 1 at the end of each block within a permuted block design) 3.unconditional allocation probabilities, being of the probabilities of all given patients randomized at the 1st, 2nd , Nth assignment to receive a given treatment unconditional on the previous assignments – these probabilities are generally constant for trials with equal allocation, but are also known to vary under several procedures with unequal allocation ratio, such as a naïve extension of the biased coin design to unequal allocation, or unequal allocation minimization (Kuznetsova & Tymofyeyev 2011), thereby potentially causing several types of bias. The only randomization for which targeted allocation probabilities, conditional allocation probabilities, and unconditional allocation probabilities coincide would be complete randomization, i.e. randomizing patients according to independent coin tosses. We would appreciate if this distinction could be made in a revised version of the guidance. Kuznetsova, O.M. and Tymofyeyev, Y. (2012), Preserving the allocation ratio at every allocation with biased coin randomization and minimization in studies with unequal allocation. Statist. Med., 31: 701-723.	We would appreciate if a distinction between targeted allocation probabilities, conditional allocation probabilities, unconditional allocation probabilities could be made in a revised version of the guidance.
EFSPI/PSI Regulatory ESIG	516	518	4.5	The guidance mentions the term "allocation scheme" in the context of adaptive randomization method. We think the term allocation scheme can be misleading here, as it could be viewed as such trials having fixed pre-generated schemes, which often is not the case as the allocations are typically generated dynamically within the IRT system. We would consider it helpful to have a definition section in order to clarify such terms.	We would appreciate if a definition could be added to the section

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFSPI/PSI Regulatory ESIG	518	521	4.5	The guidance briefly mentions covariate-adaptive randomization (CAR) in this sentence, but then drops this type of designs to only focus on response-adaptive randomization (RAR). As there are many shared aspects of CAR and RAR, like issues regarding implementation, the requirement to use randomization tests due to type I error rate inflation of frequentist methods, and also given the lack of a dedicated regulatory guidance regarding CAR, we believe this guidance should be expanded to include CAR designs.	We suggest to include CAR designs into the guideline.
EFPIA	521	521	4.51	This section only focuses on response-adaptive randomization. It could be interesting to have a recommendation about when the covariate-adaptive randomization becomes more appropriate than stratified randomization.	Please provide a recommendation for when the covariate-adaptive randomization becomes more appropriate than stratified randomization.
EFPIA	523	525	4.51	This text is appropriate in the context of the scope of confirmatory trials. There is an opportunity to mention that in non-confirmatory studies other adaptive allocation rules would be appropriate e.g. allocating subjects to doses to maximise the information about the dose-response relationship.	An example worth adding besides response-adaptive randomization (RAR) is adapatation to particpant allocation in order to optimally characterizing the dose-response relationship in an early phase trial.
EFPIA	523	525	4.51	The comment regarding the key idea is very limited and, in some cases, even misleading, as response-adaptive randomisation (RAR) is often considered a strategy to optimise performance with respect to different optimisation criteria (objectives, e.g., achieving optimal allocation), not necessarily a myopic strategy of assigning new subjects to a better-performing arm.	Remove: "The key idea is to assign new participants with greater probability to treatment arms that have had, to that point, more positive outcomes than to other treatment arms."
EFPIA	526	536	4.52	It is critical to note the additional challenges of such approaches in disease areas with substantial potential for placebo response. A well-established driver of placebo response is patient and investigator expectation of a positive outcome. Designs that are more likely to assign subjects to the 'superior performing arm' will elevate any such expectation.	An important challenge to highlight with RAR is in disease area with known placebo effect (e.g., pain). With RAR, the placebo effect may be elevated.
c4c-S	526	548	4.5	The text is difficult to read.	I would recommend to list all the challenges together in another paragraph.
EFPIA	529	548	4.52	In the presence of overall time trends any adaptive trial is at risk of bias and type I error rate inflation, not just those changing the allocation of participants. Furthermore, some non-adaptive trials could be at affected by a time trend (e.g. single arm trials). It is perhaps true that those using the latter adaptation may be at a higher risk, but the mention of time trends only in this section may be misleading as to the impact trends could have in other adaptive designs. This statement should be made in relation to general adaptive designs, perhaps in the special topics and considerations section.	Suggest adding a similar line to reflect the impact of time trends on operating characteristics in Section 5.1 (and/or in the general principles for adaptive designs).
EFPIA	529	534	4.52	The concerns regarding the use of response-adaptive randomization (RAR) detailed here assume a particularly naïve implementation of RAR that would be inconsistent with current best-practices in a setting in which a change in overall prognosis with time is plausible. It should go without saying that all clinical trial design strategies can be implemented poorly, with an attendant compromising of the validity of the trial result; it is not a valid criticism of a technique that it can be done poorly.	Line 529 could be revised to state "…valid statistical inference. If poorly or naively implemented, RAR designs…".

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFSPI/PSI Regulatory ESIG	529	531	4.5	The guidance addresses the concern that "RAR designs are susceptible to bias and inflation of the Type I error probability in the presence of overall time trends". We feel that it would be worthwhile to mention that susceptibility to bias due to time trends is a risk for all trials adapting the allocation ratio, even when these adaptations are made without taking any outcome information into account, see e.g. Altman (2018). In addition, it would be helpful to have a clear definition of what is meant by "time trend", as this term can also be viewed within other contexts, such as a diminishing treatment effect over time, which of course would also be an issue for a trial with fixed target allocation probabilities. Altman DG (2018). Avoiding bias in trials in which allocation ratio is varied. J R Soc Med. 2018 Apr;111(4):143-144.	We recomment to include a definition of "time trend".
EFSPI/PSI Regulatory ESIG	534	536	4.5	It is claimed here that RAR designs might lead to situations in which the majority of patients would be assigned to treatment arms with a high response rate but a negative benefit-risk profile. If such considerations are a concern, there are relatively easy measures to tackle this, e.g. using both efficacy and safety information to inform the allocation probabilities, as done e.g. in (Backonja et al. 2017). Backonja M, Williams L, Miao X, Katz N, Chen C. (2017) Safety and efficacy of neublastin in painful lumbosacral radiculopathy: a randomized, double-blinded, placebo-controlled phase 2 trial using Bayesian adaptive design (the SPRINT trial). Pain. Sep;158(9):1802-1812.	We think that the guideline should address the fact that there are options available for RAR designs which can tackle this issue directly and with ease.
EFPIA	543	546	4.53	Randomization ratio adaptation potentially can unveil the treatment effect even without data unblinding. For example, suppose the original randomization ratio was 1:1 for the active treatment versus placebo control. Per new requirement from regulatory authorities, to have more safety data for the active treatment, an adaptation on randomization ratio of 2:1 was made during the trial. Then the pooled sample mean before the adaptation would be approximately (mean1+mean0)/2 and the pooled sample mean after the adaptation would be (2mean1+mean0)/3. Their difference would be (mean1-mean0)/6. We can estimate the between-treatment difference even without data unblinding. Appropriate measures should be taken when a randomization ratio adaptation is performed for a trial.	Add the following sentence in line 546 "One approach that controls the Type I error probability is to allow randomization ratio adaptation at only a single or small number of interim analyses, while utilizing adaptive hypothesis testing based on pre-specified weights for combining the information across trial stages. Appropriate safeguards should be implemented when adapting randomization ratios to ensure treatment effects will not be visible."
EFSPI/PSI Regulatory ESIG	543	548	4.5	Regarding the approach to ensure type I error control, the guideline advocates to only allow adaptations of the allocation ratio only once or a few times, and to use adaptive hypothesis testing with using pre-specified weights to combine information across trial stages. Randomization tests are mentioned as an alternative, but are at the same time discouraged as it is stated that they are less powerful than a design with fixed randomization scheme. First of all, the implicit recommendation of using only one or few adaptation points can be interpreted in a way that the use of RAR methodology which continuously adapts the allocation probabilities (e.g. the Doubly Adaptive Biased Coin Design proposed by Hu & Zhang 2004) is discouraged, especially since the methodology of splitting the data by stage and using weights to combine information cannot be used for such RAR designs (while randomization tests are of course still possible to be used). We would like to request further scientific evidence for this recommendation. Similarly, the guidance claims that randomization tests "might be less powerful than a design with a fixed randomization scheme", but provides no literature evidence supporting this claim. This makes it difficult to ascertain under what circumstances which method is better suited. Furthermore, we want to draw your attention that this section stands in contradiction to what is written in lines 18-19 of the guideline, where it is stated that "[t]his guideline does not discuss the use of specific statistical methods." Hu, F. and Zhang, L-X. (2004). Asymptotic Properties of Doubly Adaptive Biased Coin Design for Multi-Treatment Clinical Trials. The Annals of Statistics 32(1) 268301.	We would appreciate if scientific evidence could be included for the recommendation. We would also appreciate if contradicting statements in the guideline could be revised.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	549	550	4.54	states that these changes should be done "without sponsor involvement" (which goes against ICH E6(R3) ' sponsor oversight)	Suggest modifying to be consistent with the ICH E6 (R3)
EFSPI/PSI Regulatory ESIG	549	550	4.5	It is mentioned that "[a]n approach that implements the changes to the randomization scheme over time without sponsor involvement should be planned to reduce the risk to trial integrity." This statement is in contradiction to ICH E6(R3), which clearly states the responsibility of the sponsor to conduct oversight of trial-related activity of service providers in section 3.9.5: "The range and extent of oversight measures should be fit for purpose and tailored to the complexity of and risks associated with the trial. The selection and oversight of investigators and service providers are fundamental features of the oversight process. Oversight by the sponsor includes quality assurance and quality control processes relating to the trial-related activities of investigators and service providers." We therefore feel that this statement should be revised and be put in the context of what is clearly spelled out in ICH E6(R3).	We would appreciate if the guideline could take into account recommendations from other guidelines - in this case ICH E6(R3).
EFPIA	553	555	4.54	Ensuring timely available high-quality interim data is essential to the integrity and validity of any adaptive trial, not just those changing the allocation of participants. This remains true even for trials with a single adaptation or a limited number of interim analyses. The pressure on data teams and infrastructures — where most trials are non-adaptive—to produce high-quality data at a different time scale than usual cannot be underestimated. This statement should be made in relation to general adaptive designs, perhaps in the special topics and considerations section.	Suggest adding this same line in the section 5.1 (and/or in the general principles for adaptive designs) and emphasise the importance of high quality timely interim data to deliver the benefits of the adaptive design. This should be part of the elements needed to deliver it.
EFSPI/PSI Regulatory ESIG	553	555	4.5	The term "randomization system" is mentioned in the context of a system that supports the randomization process of a clinical trial. We would like to point out that the randomization module is usually part of a much larger system, the so-called Interactive Response Technology (IRT) systems (which are very rarely internal systems of the sponsor, but rather systems provided by external vendors) which are fulfilling other tasks, primarily drug supply management. It might be preferrable to use a consistent term, especially because in line 865 the term "interactive voice or web randomization system", an alternative, but rather outdated term, is also mentioned as the tool that is typically used for managing randomization.	It would be appreciated if consistent language could be used throughout the guidance, preferably using the more up-to-date term "IRT".
EFSPI/PSI Regulatory ESIG	556	560	4.5	This subsection mentions the non-randomized and fully deterministic "Play-the-Winner" design proposed by Robbins (1952) and then Zelen (1969), which is clearly not RAR, as there is no random element involved. While we agree that such procedures should be discouraged, we feel that mentioning these procedures within a section on RAR could be interpreted by readers as a discouragement to use RAR in general, and an appropriate context would be helpful to prevent such misunderstandings. Robbins, H. (1952). Some Aspects of the Sequential Design of Experiments. Bulletin of the American Mathematical Society 58 527-535. Zelen, M. (1969). Play the Winner Rule and the Controlled Clinical Trial. Journal of the American Statistical Association 64 131-146.	We recommend to include more context around these methods to avoid misunderstandings.
c4c-S	556	560	4.5	The paragraph does not seem to fit here.	I would move this paragraph to line 521, after the sentence that finishes with "data". The sentences that start with "This section" to the end of the paragraph (line 525) would be in another paragraph.
ACRO	561	880	5	Challenge: While comprehensive, this section is not forward-looking enough for emerging digital and analytical methods.	Recommendation: ACRO suggests adding a discussion of machine-learning-assisted modeling and multi-protocol adaptation with the addition of subsections or examples covering:

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
ACRO	561	880	5		Machine learning–assisted response modeling and real-time data integration as permissible within pre-specified adaptive frameworks.
ACRO	561	880	5		Adaptive sub-study modifications within multi-protocol (master protocol) settings, referencing FDA 2022 and ICH E11A (model-informed drug development) for alignment.
ACRO	561	880	5		This ensures E20 anticipates future analytical and digital methodologies while maintaining rigor through pre-specification.
Cancer Research UK Clinical Trials Unit, University of Birmingham	561	561	5	There is no specific guidance on accepted approaches and considerations in the area of rare diseases. For example, the use of single arm trials may be confirmatory. It may be useful to include an additional subsection in section 5 that provides the ICH perspective on adaptive trials in rare diseases if the interpretation is more permissive or different to the content already provided.	
c4c-S	565	573	5	The requirement for an IDMC with adaptive design expertise is appropriate, but the text could clarify expectations when an independent data review committee (DRC) or hybrid oversight structure is used instead.	Add: "When a formal IDMC is not feasible, a DRC or equivalent oversight structure with comparable expertise and independence may be acceptable if justified and documented in the protocol and charter."
c4c-S	565	611	5.1	The need for independent statisticians external to sponsor may be challenging for academic trials.	Preferably external to sponsor, or functionally independent within the same institution for non-commercial trials
Takeda	567	570	5.1	An IDMC may or may not be the best candidate for making preplanned interim adaptation decision, depending on the expertise of the committee member. For complex decision making, it will be beneficial to form a dedicated adaptation committee with appropriate expertise. Similar language is available in FDA draft guidance on the use of data monitoring committees in clinical trials (Feb 2024).	If an IDMC is used for a trial with an adaptive design, it should contain, as a group, all expertise needed for making adaptation recommendations in addition to meeting its usual responsibilities (i.e., protecting individual participants' safety while maintaining trial integrity).
Cancer Research UK Clinical Trials Unit, University of Birmingham	571	571	5.1	Could the guidance incorporate an example of how the DMC statistician can be deemed to be knowledgeable and experienced as this differs to the terminology of 'trained'.	
EFPIA	572	573	5.11	It is unclear whether the statement "The IDMC should generally have access to unblinded efficacy and safety data" specificially refers to the DMC meeting in which the IDMC discussed the adaptation or to all DMC meeting. Best practice would be that the IDMC can access unblinded efficacy and safety data at each meeting.	Provide clarification when the IDMC should have access to unblinded efficacy and safety data.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EUCROF - EU CRO	572	573	5.1	"The IDMC should generally have access to unblinded efficacy and safety data."	This sentence stands in contrast to ICH E6(R3), Section 3.9.8 "To minimise bias, such committees should typically be blinded to the asigned treatments when performing their assessments, regardless of whether the trial itself is conducted in a blinded manner". It is recognised that E6 talks about endpoint assessment /adjudication committees, however, it should be clarified that in E20 - different from other IDMCs - an IDMC for a trial with an adaptive design, should have access to unblinded efficacy and safety data. Proposed Change: An IDMC for a clinical trial with adaptive design should generally have access to unblinded efficacy and safety data.
MRC Clinical Trials Unit	573	580	5.1	In row 573, the document says the IDMC should have access to unblinded data. Should this read that the IDMC should have access to unblinded analysis results? In row 580, the document says that an independent group should do the analyses. This is not consistent with row 260-263.	
c4c-S	578	584	5	The section correctly emphasizes independence, but lacks clarity on regulatory expectations for contracts, data flow, and firewalls between sponsor and unblinded team.	Add: “Sponsors should document contractual independence of the statistical group and provide a data-flow diagram showing segregation of unblinded and blinded activities.”
c4c-S	578	584	5.1	The section correctly emphasizes independence, but lacks clarity on regulatory expectations for contracts, data flow, and firewalls between sponsor and unblinded team (cf Maria (c4c-S))	Add a line cf Maria (c4c-S)'s suggestion
Ferring Pharmaceuticals	580	582		Could it be further elaborated upon why the statistical experts preparing interim reports are required to be both independent from the Sponsor and the iDMC? A statistical expert that is independent from the sponsor that has prepared the interim reports has excellent knowledge of the data and can therefore perhaps better advice the iDMC than a separate DMC statistican that only receives the interim report? Also, a scenario can be envisaged where the Sponsor prepares interim reports based on blinded data and then allows the iDMC statistician to re-run the code for the reports based on unblinded data. Could that be an acceptable set-up?	
c4c-S	585	590	5	The requirement for confidentiality procedures could be strengthened to include explicit audit trail expectations	Add: “All accesses to unblinded interim data should be logged with time, personnel identity, and purpose to allow retrospective regulatory verification.”
c4c-S	585	590	5.1	Maria (c4c-S): The requirement for confidentiality procedures could be strengthened to include explicit audit trail expectations	Line suggested by Maria (c4c-S) is OK, but also mentioned in lines 609-611 in the current version and in lines 910-915
EFPIA	587	588	5.12	While it is "strongly recommended" that unblinded results are accessible only by ISG and IDMC, the guideline later acknowledges that some degree of access to unblinded data by sponsor representative is possible (lines 602-611). Having this strong wording here can be misinterpreted.	Remove the sentence "It is strongly recommended that the independent statistical group and IDMC have sole access to unblinded interim data and results" or edit to refer to the points discussed in lines 602-611. "It is strongly recommended that the independent statistical group and IDMC have sole access to unblinded interim data and results, in certain circumstances limited sponsor access may be considered (see line 602-611)."

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	591	592	5.13	Suggest clarifying the terminology to reflect that the personnel receiving IDMC recommendations may not be directly employed by the sponsor. Specifically, we recommend changing “designated sponsor personnel” to “designated sponsor-affiliated personnel” to include individuals from CROs or other contracted organizations who act on behalf of the sponsor. This ensures the guidance accurately reflects practical trial arrangements while maintaining separation from the operational trial team.	Propose change to "Upon reviewing the unblinded interim results, the IDMC should provide adaptation recommendations to designated sponsor-affiliated personnel separated from the trial team."
EFPIA	591	592	5.13	Not all adaptations require IDMC review and approval, e.g., routine updates to RAR proportions within specified bounds, and this possibility should be acknowledged within the draft Guideline.	Please consider adding, between lines 590 and 591, a new paragraph to read: “Many, but not all, adaptations to be implemented based on prespecified decision rules should be reviewed by an IDMC or similar body prior to implementation, to ensure that the prespecified rule remains scientifically and ethically appropriate. In this context, the IDMC should be aware that deviations from the prespecified rules may compromise the integrity and operating characteristics of the trial, so should only occur when necessary. There may be some more routine adaptations, e.g., routine updates to RAR proportions within specified bounds, that do not require IDMC review prior to implementation.”
EUCROF - EU CRO	591	601	5.1	In this section, the sponsor is addressed several times. Whereas in line 592 it is clear that sponsor representatives seperate from the trial team are meant, it is not totally clear lateron when the sponsor is mentioned and should be described in an unambiguous way.	Proposed changes: In the specific case that the IDMC has made a recommendation to stop a trial early, sufficient information may then be communicated to the sponsor (e.g., key efficacy and safety results) to allow sponsor representatives independent of the trial team decision-making about whether to stop the trial. In general, however, the adaptations should be planned such that the sponsor trial team can implement the IDMC recommendations regarding trial adaptations without having access to any unblinded interim results.
c4c-S	591	601	5.1	This paragraph should have a subtitle.	I would call it "procedures"
c4c-S	591	599	5	The draft provides examples of IDMC-to-sponsor communication but omits expectations for adaptation documentation and contemporaneous record-keeping	Add: “Adaptation decisions and rationale should be documented in contemporaneous records, retained in the Trial Master File, and made available for regulatory review.”
c4c-S)	591	599	5.1	The draft provides examples of IDMC-to-sponsor communication but omits expectations for adaptation documentation and contemporaneous record-keeping (cf Maria (c4c-S))	Add a line cf Maria (c4c-S)'s suggestion

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFSPI/PSI Regulatory ESIG	595	597	5.1	It is mentioned that “the adaptations should be planned such that the sponsor can implement the IDMC recommendations regarding trial adaptations without having access to any unblinded interim results”. We consider this recommendation to be very difficult to implement in practice, as adaptations will inevitably require implementation within IRT systems. These may either be with the sponsor itself, or will need to involve the sponsor for drug supply, shipment, and randomization schedule. Therefore, we appreciate some critical review of this recommendation, taking the mentioned issues in implementing such an approach into account.	We suggest that the guideline should acknowledge the fact that implementing adaptations without sponsor involvement is operationally challenging. Adaptations often require updates in IRT systems for drug supply, shipment, and randomization, which typically involve sponsor participation. Instead of excluding the sponsor entirely, we propose emphasizing robust confidentiality measures (e.g., role-based access controls, use of internally unblinded functions not involved in trial activities) to prevent disclosure of unblinded data while allowing practical implementation of IDMC recommendations
EFPIA	602	605	5.14	Examples of specific situations, where this would be acceptable, would be welcome.	Provide example of situations, where unblinded access would be acceptable and to whom - or delete "However, sponsors may propose some degree of access to unblinded data in certain circumstances”
EFPIA	602	611	5.14	Can the guideline provide examples of the "certain circumstances" where it is acceptable to have some restricted access to unblinded data by the sponsor. This could include cases such as treatment selection of a drug or a dose, which has major impact for the sponsor, and final decision should be made by the sponsor, based on recommendations from IDMC, and after review of unblinded data. Examples could includes cases where the adaptation recommendations can have major implications for the study/project/company, such as stopping a study for futility, or for overwhelming efficacy leading to early submission and registration.	[proposal] "However, sponsors may propose some degree of access to unblinded data in certain circumstances, particularly when adaptation recommendations have major implications for the study, such as stopping the trial for futility or overwhelming efficacy that could lead to early submission and registration.This should be made explicit at the planning stage."
Inmaculada Baeza (c4c-S)	602	611	5.1	The paragraph does not seem to fit here.	I would move this paragraph to section 3.5
c4c-S	602	610	6	The section appropriately restricts sponsor access but could benefit from guidance on exceptional access (e.g., safety signals or supply chain modifications).	Insert: “Sponsor access may be permitted in exceptional circumstances (e.g., urgent safety concerns or major logistical adaptations), provided justification, independence, and full documentation are ensured.”
EFPIA	603	604	5.14	At lines 848-849 (Section 5.5 Adaptive Designs in Exploratory Trials), it is acknowledged that dose selection decisions "may entail multidimensional adaptation decisions that require considerable input from various disciplines within the sponsor". This challenge is not unique to exploratory dose-ranging trials; it may equally apply when dose selection is based on an interim analysis in confirmatory trials. Therefore, it is suggested to explicitly mention this scenario as an example of a circumstance in which it may be reasonable for the sponsor to propose a certain degree of access to unblinded interim data.	However, sponsors may propose some degree of access to unblinded data in certain circumstances (for example, in case of complex dose selection decisions).
EFSPI/PSI Regulatory ESIG	603	604	5.1	"However, sponsors may propose some degree of access to unblinded data in certain circumstances." Examples for such circumstances may help guide discussions in cross-disciplinary teams with different views.	Please consider adding examples for circumstances where sponsor access to unblinded data is considered acceptable to some degree.
EFPIA	612	725	5.2	The section gives the impression that there is a clear implicit recommendation on the statistical paradigm (i.e., frequentist) in which adaptive designs should be planned or analyzed. This is confusing because other sections seem to go lengths in avoiding such an implicit recommendation.	Remove reference to any operating characteristic that is specific to statistical paradigms (e.g., frequentist, Bayesian, etc) or be very explicit early in the document that there is a clear expectation on the statistical paradigm.
c4c-S	612	725	5.2	No comments	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	613	629	5.21	An important use of simulation which is missing from this overview is the ability to use simulations to quantify bias in estimation of treatment effects and, specifically, to determine whether the bias, if any, is of a sufficient magnitude to require the use of alternative estimation methodology or the alteration of a design feature, e.g., the timing of a first interim analysis. This use is mentioned briefly on line 638 and in more detail on lines 646-647; however, it should be included in the introductory summary of the uses of simulation because of its importance.	Please consider adding the following text on line 625, between the two existing sentences: "An important use of simulations to quantify bias in estimation of treatment effects and, specifically, to determine whether the bias, if any, is of a sufficient magnitude to require the use of alternative estimation methodology or the alteration of a design feature, e.g., the timing of a first interim analysis."
c4c-S	613	625	6	Simulation Principles and Usage	Require version-controlled simulation code with reproducibility checks and sensitivity analyses on Type I error control.
EFPIA	614	620	5.21	It may not be necessary to limit the statement to operating characteristics alone, as simulations can also be used to assess estimates along with their potential biases, variabilities, study duration, sample size and other relevant factors.	Suggest to change "oprerating characteristics" to "operating characteristics and other important properties"
EFSPI/PSI Regulatory ESIG	618	618	5.2	The draft guidance mentions the term "dropout rate" multiple times. Perhaps also the impact of varying assumptions regarding the frequency of post-baseline events impacting the interpretation or existence of the outcome of interest could be mentioned?	Please consider pointing out that in addition to dropout (mostly leading to missing data) also varying assumptions on the frequency of post-baseline events impacting the interpretation or existence of the outcome of interest may need to be accounted for?
c4c-S	626	642	6	Design Comparison via Simulation	include non-adaptive benchmark simulation outputs in regulatory submissions to contextualize performance.
Inmaculada Baeza (c4c-S)	630	642	5.2	The paragraph should have a title/subsection/subheading	
EFPIA	634	635	5.22	A key point of clarification is needed to prevent a potential misunderstanding within the guidance document. The current text could be read as somewhat equating group sequential designs (GSDs) with conventional, non-adaptive designs. This would be misleading as, by the very definition provided in the guidance, a GSD is an adaptive design where the sample size can be altered based on pre-planned stopping rules. The impression that GSDs are non-adaptive is unhelpful for two reasons. Firstly, it creates a contradictory and confusing message for users of the guidance. If a GSD is not considered an adaptive design, it undermines the very purpose of a guidance document on adaptive trials. Secondly, the use of "well-understood" echoes the draft version of the U.S. Food and Drug Administration (FDA) guidance on adaptive designs, and this wording arguably created two 'classes' of adaptive designs that (unintentionally) penalised the use of those deemed as "less well-understood". We recommend a change to the wording here to ensure that GSDs are clearly positioned within the adaptive design framework, as their well-understood operating characteristics make them an excellent starting point for those new to adaptive methodologies.	Suggest editing the text here to say, "These should include a well justified benchmark design and analysis approach which could be a non-adaptive design but not necessarily so."
EFSPI/PSI Regulatory ESIG	634	634	5.2	"well-understood" appears to be in the eye of the beholder. It means something else for a student compared to someone who has worked in drug development for 30+ years.	Replace "well-understood" by "transparent"
EFPIA	635	636	5.22	Could clarifications on whether the requirements stated for reporting the operating characteristics of the design (e.g. number of simulation runs and nuisance parameters etc) pertains only to the chosen design or if there is an expectation to see all the evaluated designs	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	635	642	5.22	The preceding lines discuss the need for simulation to adequately explore the chosen design and compare its performance to relevant, potentially simpler designs. However, the text on lines 635-642 could be interpreted as expanding this discussion to suggest including in regulatory submission the performance of other designs of potentially equal complexity. While we wholeheartedly endorse such simulations in the planning phase of a trial, we would recommend excluding such non-selected designs in a submission to regulators. The selection of a design among equally complex options is typically governed by sponsor-specific criteria (cost of interim analyses versus efficiency, expected time to completion relative to the competitive landscape, etc.). If the selected design meets the criteria outlined in the document and is superior to alternative, simpler designs, it is substantially and overly burdensome to both prepare and review the full, often months long process of design selection among equally complex options.	We suggest removing all text suggesting a need to submit simulation results for other designs of potentially equal (or greater) complexity that are not being proposed.
EFPIA	643	648	5.23	When listing key operating characteristics to be assessed in the simulated studies, only expected sample size is listed. It might be helpful also to include minimal sample size (related to benefit-risk assessment) and max sample size (related to affordability and meaningfulness when a weaker treatment effect is claimed to be statistically significant). However, Lines 650-655 also discuss aspects of key operating characteristics, and it is unclear what is specifically covered. Please clarify.	Please consider adding information / detail as suggested. Please also clarify lines 650-655
c4c-S	643	655	5.2	The paragraph should have a title/subsection/subheading	
EFPIA	644	654	5.23	Listing frequentist measures such as Type I error probability, coverage of confidence internals, etc as mandatory operating characteristic that a simulation study must include is not sensible unless the aim is to exclude Bayesian designs.	Replace 'type I error probability' with 'probability of erroneous conclusions'; remove reference to frequentist operating characteristics
MRC Clinical Trials Unit	645	645	5.2	There are various forms of type 1 error (e.g. average type 1 error, doi:10.1080/19466315.2024.2342817). This should be left flexible.	Change to "measures of type I error probability".
MRC Clinical Trials Unit	646	646	5.2	I recommend clarifying the discussion on bias to avoid potential misinterpretation. Bias can be defined in different ways (e.g., precision-weighted bias, conditional bias) and its meaning depends on the specific scenario. To enhance clarity, any reported bias should ideally be accompanied by the corresponding probabilities (such as the probability of stopping or continuing). Providing these probabilities offers valuable context and allows the reported biases to be interpreted in terms of their likelihood of occurrence, which is essential for informed decision-making.	Change to "measures of bias". Add probabilities of stopping and continuing.
c4c-S	651	655	6	Summary Metrics	Suggest requiring presentation of variability (e.g., interquartile ranges) alongside averages to prevent misinterpretation of mean-based summaries
EFPIA	656	672	5.24	There is a need to be more explicit about the range of treatment effects evaluated. It is important to state that simulating two scenarios (e.g., the Null and the assumed effect used in the sample size calculation) is inadequate. It may also mention that intermediate effects are why "characterising the p value distribution" is preferable to solely relying on error rates.	Propose updated language stating that a a range of treatment effects should be covered, rather than just the null and the assumed treatment effect for sample size determination.
EFPIA	656	656	5.24	"the plausile range of assumptions" is too narrow as "plausibility" is too subjective a concept. In particular, as said in the "recommendation" section, we usually would have to investigate null scenarios of no effect of the experimental treatment, even if these would be deemed „implausible" by many experts. In general, it is too easy to resort to „plausibility" when dismissing secnarios that yield undesired results. It has to be made clear that the range of scenarios must be comprehensive.	Replace with "a wide range of assumptions comprising a variety of plausible assumptions, but also potentially some null scenarios even if these are deemed implausible."
c4c-S	656	679	5.2	The paragraph should have a title/subsection/subheading	

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Invents consortium - EU Horizon project	656	679	5.2		Settings where multiple hypotheses are tested should be explicitly discussed and it should be stated that not only the global null hypothesis but all combinations of true and false null hypotheses (with a range of effect sizes for the latter ones) need to be investigated to demonstrate type 1 error rate control. (see also the comment on lines 374-378) on the definition of the term "nuisance parameter".
EFPIA	660	660	5.24	Please clarify the meaning of 'greater' control? Does it mean the faster enrollment or lower dropout rate? or both?	
EFPIA	660	664	5.24	Simulations should not be conducted under all plausible scenarios and on fine grids. Like in standard experimental design practice, simulations should be conducted for relevant scenarios to learn on the question of interest. Adequacy of assumptions appear to be driven here by clinical/statistical considerations. It may be worthwhile to consider simulations as a type of statistical experiment, which can be actively designed to understand operating characteristics. As conventional in design of experiments, the design settings would then be chosen to learn as much as possible on the question of interest. A "fine grid" may then not be required.	Revise the section with respect to the selection of scenarios for simulation.
EFPIA	665	670	5.24	Please provide recommended metrics & examples how fine the grid should be.	
Invents consortium - EU Horizon project	668	668	5.2		If a worst case scenario with regard to the type 1 error rate can be identified, the number of simulations scenarios can be reduced considerably. It seems that this is alluded to by the reference to "monotonicity arguments"? This should be more clearly stated.
c4c-S	671	679	6	Uncertainty in Type I Error Estimation	Suggest emphasizing regulatory expectation for larger simulation replicates when adaptive decision rules are complex or involve multiple adaptations
EFPIA	677	678	5.25	For many adaptive designs that require simulation for determination of type I error risk, the direction of the effects of nuisance parameters and other factors on type I error is easily known. This allows the determination of the "worst case" type I error risk within the plausible range of these parameters. Thus, the "additional uncertainty" mentioned here may or may not exist and it would be more accurate to write "Thus, there may be additional uncertainty for designs…" and "When additional uncertainty exists, additional justification…".	Please consider the suggested change.
c4c-S	680	691	5.2	The paragraph should have a title/subsection/subheading	
EFPIA	686	688	5.26	It seems this could be challenging to conduct 100000 simulations if computationally intensive modeling involving individual patient data (Bayesian or otherwise) and multiple scenarios are needed. In the Bayesian case if methods are known not to control type one error, but would limit erroneous conclusions, would 100000 simulation be needed for a more general understanding of error control? The document should better clarify this potentially in section 5.3	Suggest clarification of the requirements in cases where the focus of error control is not type one error potentially in section 5.3
EFPIA	686	687	5.26	It is wortwhile to add here that precision of simulation quantities of interest should be monitred with the monte carlo standard error. The number of replicates should be chosen in line with a desirable precision.	Suggest reporting of MCSE as a tool to document precision.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	686	687	5.26	The simulations should be designed to minimize noise. For example, when the data generating process is not varied, but the rules for adaption, then these comparisons should be done on the same simulated trial populations allowing for paired comparisons and therby reducing noise.	Encourage reduction of simulation noise as discussed in the reference doi:10.1093/biomtc/ujaf012
EFPIA	687	687	5.26	If MCMC is used, Bayesian simulation takes so long and it may not be feasible to iterate 100,000 times. If there is suggestion for number of iterations to simulate Bayesian study, it is helpful.	Replace "For example, " with "For example and if feasible, "
EFPIA	687	687	5.26	We suggest to not require a specfic number of simulation (100.00 is stated) since the computing power is under rapid development and the number should be chosen and justified based on the complexity and critically of the specific suimulation	
EFPIA	687	688	5.26	The accuracy of operating characteristics directly depends on the number of repetitions. Since a simulation-based operating characteristic value will always be random and may be different from the true/targeted value.	Please provide the maximum acceptable error margin a key operating characteristic. E.g., type I error must be controlled at alpha +/- 0.1%
c4c-S	692	725	5.2	The paragraph should have a title/subsection/subheading	
c4c-S	692	725	6	Reporting and Documentation of Simulations	Suggest numbering items 1–11 as mandatory checklist elements for submission, and encouraging visual dashboards (e.g., interactive Shiny tools) for regulators' interpretation.
EFSPI/PSI Regulatory ESIG	693	694	5.2	"regulatory submissions prior to conducting the trial" Generally, it is enough to implement the final design prior to the first cutoff for an interim analysis.	Change to "regulatory submissions prior to conducting the trial prior to the first clinical cutoff for an interim analysis"
Invents consortium - EU Horizon project	698		5.2		The range of futility boundaries considered in simulation studies should also be specified, particularly if a non-binding futility is proposed when multiple futility rules could be considered
EFPIA	712	713	5.27	Could a clarification be provided as to what is meant by an interactive graph and how it is intented to be included in a simulation report, which seems to be understood as a document?	
EFPIA	714	715	5.27	Please provide guidance how to select a "representative" iteration from a large number of simulations.	
EFPIA	718	718	5.27	Presenting individual trial results from a simulation also facilitates understanding of the *decision rules* at design stage (which may not be straightforward to understand for a non-statistician)	Suggest revising "better understanding of interim decision rules and potential interim decisions..."
EFPIA	721	722	5.27	Please provide an example to ilustrate.	
EFPIA	721	722	5.27	It is unclear what clinical discussion could be made about simulation results. Suggest this is removed from guidance or additional guidance on scope provided.	[proposal] Remove the bullet point "11. A clinical discussion about if and to what extent the simulation results address the key questions." or provide additional clarification on the specific nature and scope of clinical discussion expected regarding the simulation results."

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	723	725	5.28	It is unclear whether this applies to any adaptive design, i.e. also for simple GSD or SSR designs.	Clarify if the statement also applies to GSD and SSR designs.
EFSPI/PSI Regulatory ESIG	723	725	5.2	"The careful documentation of simulation studies is also critical because the validity of the simulations and associated conclusions will be part of the regulatory review of results at the end of the trial." We feel it is late to only have clarity about conclusions from simulations at the time of review. We therefore suggest to discuss this earlier in the process.	Change to "The careful documentation of simulation studies is also critical because the validity of the simulations and associated conclusions will be part of the trial design discussions."
ACRO	726	824	5	Challenge: ACRO supports the inclusion of Bayesian and time-to-event methods, but this would benefit from further elaboration.	Recommendation:
ACRO	726	785	5.3	Challenge: ACRO notes that much of discussion in this section is focused on prior distribution considerations that are not specific to adaptive designs.	Recommendation: ACRO suggests providing guidance on reporting of Bayesian results, including justifying sensitivity analyses for selected prior choices and handling posterior probability thresholds for decision-making.
EFPIA	726	726	5.3	The footnote states that "This section on Bayesian methods for adaptive designs is not fully harmonized." Considering this is an ICH guidance and that the purpose of ICH is to foster the adoption of global standards, it is crucial that a resolution is found for a globally acceptable approach to Bayesian methodologies. The use of Bayesian methods for external data borrowing is particularly disincentivized in this guidance, with a requirement to evaluate the trial data with no borrowing, which is a particularly stringent requirement compared to other adaptive approaches which may be seen as disproportionate; in the future, it would be good to have trust in these approaches.	The use of Bayesian methods for external data borrowing is particularly disincentivized in this guidance, with a requirement to evaluate the trial data with no borrowing, which is a particularly stringent requirement compared to other adaptive approaches which may be seen as disproportionate; in the future, it would be good to have trust in these approaches.
EFPIA	726	726	5.3	The use of Bayesian methods with data borrowing may be justified beyond the scope of rare diseases and pediatric clinical development when there are ethical benefits. Instead of exposing participants to sub-optimal treatment options, it could be envisaged to use data borrowing considering the data is already available for this treatment in this indication. Sponsors and health authorities should base these decisions on the expected benefit of the treatment relative to the potential risk.	The use of Bayesian methods with data borrowing may be justified beyond the scope of rare diseases and pediatric clinical development when there are ethical benefits. Instead of exposing participants to sub-optimal treatment options, the guideline should indicate that it could be envisaged to use data borrowing considering the data is already available for this treatment in this indication.
EFPIA	726	785	5.3	The header concerns bayesian methods in general but only two specific use cases of bayesian methods are adressed. The case where bayesian methods are used for interim analyses and the final analysis is frequentist and then the case where bayesian methods are used for borrowing data from another trial. While these are important use cases then there many more and guidance on general usage of bayesian methods are needed	
EFPIA	726	785	5.3	Too little discussion on the "maximum weight" of external data.	Add a paragraph on approaches to avoid that the historical data "swamps" the data generated in the trial.
c4c-S	726	747	7	Bayesian Adaptive Designs	Recommend adding cross-reference to ICH E9(R1) on estimands and sensitivity to priors. Emphasize that Bayesian adaptations must still ensure Type I error control when used for confirmatory purposes.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Regeneron Pharmceuticals, Inc.	726	785	5.3	Based on the definition of adaptive design, Bayesian information borrowing from external data sources is not an adaptive design if it is only used in the final analysis, even if the extent of borrowing is dynamic and may depend on the trial data. It is however possible to use Bayesian methods to adapt the trial in terms of other design parameters such as sample size. Regeneron recommends that this point be clarified in the guidance.	
c4c-S	726	785	5.3	Bayesian methodology: no commentsx	
Cancer Research UK Clinical Trials Unit, University of Birmingham	731	731	5.3	The word potentially is superfluous as Bayesian methods are applicable	Remove 'potentially'
EFPIA	731	732	5.31	"Potentially applicable" seems inappropriate. Either they are applicable to a variety of adaptive designs or not. As examples are presented later, it may be better to describe that they are applicable, but not for all purpose and in all situations.	"Bayesian approaches are applicable to a variety of adaptive designs, even though not all adaptive designs require Bayesian approaches for analysis and implementation".
EFPIA	732	733	5.31	The current Section 5.3 provides very ambiguous guidance and will cause confusion. The statement that the principles laid out in Section 3 (trial specific type 1 error control (see line 172) and unbiasedness) should be followed, rules out Bayesian analyses that utilize external data. It remains unclear then why this section is so long and provides specific guidance on Bayesian approaches that utilize external data (which then won't fulfill the laid out principles).	Either (i) amend principles to include Bayesian inference (i.e. information combination by the formal rule of Bayes theorem) as an alternative inference mode to standard trial-specific frequentist statistics or (ii) shorten section 5.3
BSWG	736	737	5.3	The term: 'false positive conclusion' is probably more appropriate in general, and specifically in the context of Bayesian adaptive designs, instead of Type-I error. Suggest changing 'Type-I error probability' terms to 'false positive rate' (as already done in Section 5.2)	
Cancer Research UK Clinical Trials Unit, University of Birmingham	736	737	5.3	Is type I error a relevant parameter within the Bayesian framework? This sentence could benefit from clarification.	
EFPIA	736	737	5.32	The guideline provides an example of a study where the adaptation rule is Bayesian but the final analysis of the study is a frequentist analysis that controls type I error. However at the end of the paragraph, it's unclear why a Bayesian study should have control of the type I error rate.	Recommend to clarify the relationship between the Bayesian adaptation rule and the Frequentist final analysis reduce confusion on needed operating characteristics.
EFPIA	736	747	5.32	Increased clarity is required in differentiating between designs, which use Bayesian approaches for borrowing information in the primary analyses and Bayesian approaches, which are just used to inform adaptation decisions. This section seems targeting the use of Bayesian approaches to inform adaptation decisions. Still, in reading the exact focus becomes a bit unclear.	
EFPIA	736	759	5.32	Only when the historical data show a positive treatment effect, we borrow historical information in a Bayesian design. In that case, no matter how much historical data we borrow, the Type I error probability will be inflated. Please provide details on how to evaluate the Type I error rate when a Bayesian method is applied.	In line 747 provide an additional sentence to specify how to evaluate the Type I error rate when a Bayesian method is applied.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	736	737	5.32	Bayesian methods simply cannot control Type I error. This language should be adapted. Errors rates can be demonstrated to be within certain ranges given the design of the trial. The use of trial external data or other prior information only vanishes if the trial data dominates in relative size, but in that case the use of Bayesian techniques is commonly questionable.	Replace "Type I error probability is controlled" with "error rates as Type I error are reasonable for the trial design". The word controlled should really be dropped. This language should be consistently changed whenever Bayesian methods are referred to - in that case any error rate must be evaluated in the context of the trial design (after all Bayes conditions on the trial data)
Invents consortium - EU Horizon project	736	737	5.3	It is stated that "Bayesian methods can be used to inform adaptations in a trial where decision criteria for the primary analysis are chosen to ensure that the Type I error probability is controlled". Bayesian inference is not considering the Type I error which is based on frequentist hypothesis testing. This statement is missleading an can induce to error when considering Bayesian inference.	Bayesian methods can be used to inform adaptations based on posterior or predictive probabilities. It can help quantify the probability of wrong decision according to a predefined clinically based thereshold
Invents consortium - EU Horizon project	736	747	5.3	The entire section is wrongly mixing frequentist hypothesis testing concept with Bayesian inference based decision making. The way it is written is missleading, one does not depend on the other.	The entire section should be re-written using Bayesian correct methodology
Invents consortium - EU Horizon project	740	740	5.3	It is stated that "predictive probability of rejecting the null hypothesis at trial". Predictive probabilities are not constructed in there meaning to reject the null hypothesis.	This should be deleted
Cancer Research UK Clinical Trials Unit, University of Birmingham	746	747	5.3	Is type I error a relevant parameter within the Bayesian framework? This sentence could benefit from clarification.	
Ferring Pharmaceuticals	746	747		Consider adding at the end of the sentence "and ensuring an appropriate bias-variance trade-off for treatment effect estimates". The current sentence puts a lot of focus on type I error control, whereas reliable estimation is perhaps equally important?	Add at the end of the sentence "and ensuring an appropriate bias-variance trade-off for treatment effect estimates"
EFPIA	748	756	5.33	Rather than feasibility, having reliable and relevant data should be the key consideration for adopting a Bayesian analysis	Propose revising the text to emphasize reliable and relevant data as the key consdieration when adopting a Bayesian analysis
Invents consortium - EU Horizon project	748	785	5.3	Bayesian Borrowing seems out of scope for this guidance. Bayesian methods can serve distinct goals. One of them is to facilitate an easier statistical analysis. Especially for intricate adaptive designs, the Bayesian statistical analysis may be substantially easier than the frequentist counterpart. Another goal may be to include external information, for instance via Bayesain borrowing methods. The latter, in my opinion, is a fully separate goal from making a trial adaptive. We borrow when it is impossible for the individual experiment to contain sufficient information, for instance in extremely rare diseases or for paediatrics. In that case, it is unreasonable to demand type-I error control. If this were possible, we would not need to be borrowing to begin with. For paediatric extrapolation, we also do not attempt to control type-I error. We acknowledge that it is impossible to get sufficient information with the trial and try to fill this gap with external sources, potentially in a Bayesian way. This is fully distinct goal from adding flexibility to a design. Ofcoure, adaptive designs may still use Bayesian analyses because those are more practical. In that case, the first goal I describe, it is absolutely logical to demand type-I error control. We should not have a lower bar just because we chose a Bayesian analysis. Bayesian borrowing, however, does not have as a goal to add flexibility to the trial, but to include external information. There is a lack of guidance on Bayesian techniques, such as this one, but I do not think this E20 is the place to discuss Bayesian Borrowing.	Make the distinction between Bayesian methods that allow for a flexible design and Bayesian methods that aim to include external information because it is impossible to obtain sufficient information with the trial results alone (e.g., very rare diseases). In the former, type-I error control is an entirely reasonable demand; the bar should not be lowered depending on the analysis of choice. In the latter, this demand makes the whole idea of Bayesian borrowing pointless. There can, in general, be no efficiency gains if type-I error needs to be (stricly) controlled. The latter of the two goals, including external information, is not inherent to adaptive designs and should in my opinion not be in this guidance.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Invents consortium - EU Horizon project	748	785	5.3		The inclusion of external data, as, e.g., external controls in RCT has many more aspects than Bayesian borrowing (see e.g. Burger, Hans Ulrich, et al. "The use of external controls: To what extent can it currently be recommended?." Pharmaceutical Statistics 20.6 (2021): 1002-1016.) and it seems out of scope of this guidance to cover all these aspects. Regarding analysis methods, for example, statistical (causal inference) methods to adjust for confounding will be essential (in addition to dynamic Bayesian approaches to address remaining biases)
Institute of Medical Biometry (IMBI) Heidelberg, Germany	748	759	5.3	This paragraph discusses the use of Bayesian methods to borrow external information. In particular, it warns that "[m]isspecification of the prior distribution can lead to lack of control of the probability of false positive conclusions". In the most common analysis cases, namely in the presence of uniformly most powerful tests (UMP tests), power gains are not possible when control of the Type I error probability is required (see Kopp-Schneider, Calderazzo, Wiesenfarth. Biom J. 2020 Mar;62(2):361-374. doi:10.1002/bimj.201800395). For this reason, the question arises in what situation Bayesian borrowing of external information would be acceptable. If on the one hand, Type I error control is strictly required in all cases, any use of Bayesian borrowing will require a thorough discussion of the trade-off between increase in Type I error and the putative benefit of borrowing (e.g. power gain). If however, there may be situations in which deviation from the Type I error control principle is justified, guidance concerning the nature of these situations would be beneficial in combination with a discussion of what other quality measures should be used instead (e.g. average Type I error across a prior distribution of possible scenarios).	
c4c-S	749	781		In paediatric settings, Bayesian borrowing from high-quality adult data or historical paediatric registries may ethically reduce placebo exposure and sample size requirements. Guidance should highlight safeguards against over-borrowing that might obscure developmental differences in efficacy or safety.	
c4c-S	749	784	7	Borrowing of External Data	Suggest requiring explicit pre-specification of maximum borrowing weight and conflict thresholds (e.g., robust mixture priors) to ensure transparency.
PhaseV Trials, Inc.	751	755	5.3	While the use of Bayesian methods to borrow from historical data is discussed, no examples of when this might be justified are provided. Bayesian methods have long found application in rare and pediatric disease, where traditional statistical methods are known to be infeasible, yet there exist reliable historical data to assist.	Suggest adding the sentence, "Rare and pediatric disease research offers an example of a setting where Bayesian approaches of this type can be justified."
EFPIA	755	755	5.33	Misspectification of the prior is not a clear term. To some degree, one may assume many prior distributions to be misspecified, in particular if distributions are calibrated to provide adequate operating characteristics. Therefore it is recommended to adjust the language here.	Replace "Misspecification of the prior distribution can lead to lack of control of the probability of false positive conclusions" with "The Bayesian model should be calibrated to control the probability of false positive conclusions at an acceptable level".
EFPIA	755	755	5.33	"Misspecification of the prior distribution...". This sounds like a "true" prior exists, which is misleading, a prior distribution always contains a subjective element.	
EFSPI/PSI Regulatory ESIG	755	755	5.3	"misspecification" Is mis-specification really the problem? Kopp-Schneider, Calderazzo & Wiesenfarth (2020) have shown that an informative effect size prior with information favoring the experimental treatment always leads to Type-I error inflation, and if Type-I error is held constant a Bayesian design does have higher power than the corresponding non-Bayesian design.	Replace "misspecification" by "Informative priors"

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
ACRO	756	759	5.3	Challenge: This sentence claims ensuring a “Bayesian prior” reflects relevant available information introduces additional uncertainties beyond those associated with frequentist inference. However, this statement is incomplete. An uncertainty in frequentist inference, but not Bayesian analysis (which is rarely discussed) is what does a probability about data given a hypothesis tell us about the hypothesis itself? For example, a p-value is a statement about possible data that could come from a hypothetical repetition of the clinical trial, assuming the tested hypothesis (e.g. the hypothesis of no effect). Therefore, when a p-value is reported in a trial, what does it tell us about that hypothesis (leaving aside the more fundamental question of whether that hypothesis is of interest in the first place)? Outside of the Bayesian paradigm, the latter uncertainty is more difficult to overcome.	Recommendation:
ACRO	756	759	5.3		We suggest the following refined language for the final guideline:
EFPIA	756	756	5.33	If "control of false positive conclusions" (within a trial) is the guiding principle for choosing prior distributions, only priors with very limited information content would satisfy this requirement (e.g. usage of external data would be ruled out).	
EUCROF - EU CRO	756	776	5.3	In lines 756,761, 773, 774, 775, 776 (for example), the term "prior" is used. We think, in order to increase readability, "prior" should be replaced by "prior probability distribution" or "prior distribution".	**Proposed Change:** See column on the left.
EFPIA	758	759	5.33	It is not necessary to compare the Bayesian approach to a Frequentist approach here. Simply drop this statement "that are not present when using frequentist inference with no borrowing". There are other issues with Frequentist approaches in the described setting (as likely leading to inconclusive results).	Recommend to drop comparisons to a Frequentist procedure.
EFPIA	758	759	5.33	Borrowing of information and the expected reduction in expected mean-square error associated with some approaches are not unique to Bayesian estimation. The improvement in the mean-square error with James Stein estimation is long established in a frequentist context or, for example, with the use of frequentist hierarchical random-effects models. The point made here, that borrowing information that is not “fit for purpose,” e.g., is not representative of the likely true treatment effect will increase uncertainty or bias in estimated treatment effects, is true in both Bayesian and frequentist contexts. Both statistical paradigms are vulnerable to non-representative data whether those data are analyzed alone or incorporated indirectly through borrowing.	Please consider rewording the end of the paragraph, beginning on line 756, to read: “Ensuring that a prior accurately reflects complete and relevant available information is critical to ensuring valid inference.”
EFPIA	766	770	5.34	In the context of external data the text mentions that data from randomized controlled trials and recent data are generally preferred and patient-level data are generally expected We agree that having this is desirable. If 748-756 is adopted and feasibility is a key factor, there should be more flexible language here. Acknowledging that it might not be feasible to obtain external data from a recent randomized trial	Suggest text is more flexible acknowledging that external data from recent randomized trials might not be possible to obtain when a randomized trial is challenging due to feasibility and other relevant and reliable sources should be consdiered
EFPIA	767	768	5.34	While we agree that patient level data are preferred, many Bayesian borrowing methods such as power prior and meta-analytical-predictive prior only require summary level data, if no patient-level covariate effect is of interest.	Acknowledge that the use of summary level data only is acceptable for certain methods.
EFPIA	767	770	5.34	Patient-level data are rarely available to build a prior distribution.	"If available, use of patient level data allows a thorough evaluation…"

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	767	770	5.34	The statement that "(p)atient-level data are generally expected" when external information is used is overly simplistic and limiting. As noted, when such data are readily available, they may be of tremendous value; however, the advantages of using aggregated data for which, e.g., only summary statistics are available, may outweigh the disadvantages. In many cases, such data are down-weighted, and this partially mitigates the risks associated with the inability to adjust for patient-level covariates in the analysis.	Please consider revising the end of the paragraph, beginning on line 767, to read: "Patient-level data, if available, are generally of the greatest value because they allow a thorough evaluation at the planning stage of the relevance of the external information and may facilitate strategies to address potential conflict between the prior and current trial data at the assessment stage. However, using aggregated data, e.g., for which only summary statistics are available, may also be advantageous compared to omitting relevant external information altogether. In many cases, such data are down-weighted, to mitigate the risks associated with the inability to adjust for patient-level covariates in the analysis."
EFSPI/PSI Regulatory ESIG	767	767	5.3	"Patient-level data are generally expected because they allow a thorough evaluation at the planning stage of the relevance of the external information and may facilitate strategies to address potential conflict between the prior and current trial data at the assessment stage. " This appears overly restrictive. Generation of synthetic control data to overcome privacy challenges would then not be an option.	Remove the sentence entirely.
EFSPI/PSI Regulatory ESIG	767	770	5.3	"Patient-level data are generally expected ...". This sentence raised some confusion.	Perhaps consider adding some clarification? For what purpose would patient data on the level of the individual be needed and for what purposes would distributional data suffice?
Invents consortium - EU Horizon project	767	767	5.3		In addition to data from recent trials, prospective external data should be preferred. This can partially address the problem of lack of pre-specification which is difficult to address if historic data already in the public domain is used (as, e.g., clinical trial data whose aggregate results are published).
PhaseV Trials, Inc.	771	772	5.3	While it is certainly necessary to explain the degree of borrowing in any given application, I don't think the precise amount of borrowing necessarily needs to be prespecified, as we may not know this amount a priori. Rather, it is the *algorithm* for how the borrowing will be done that must be prespecified. A wide variety of cautious yet data-driven methods for adaptive borrowing from historical data have emerged over the past 10 years, including power priors, commensurate priors, robust mixture priors, elastic priors, and so on. FDA employees have even served as coauthors of many of these papers, including variants that use propensity matching to provide further protection against borrowing that turns out to be unwarranted.	Suggest changing "including the amount of borrowing from the external data" to "including the precise algorithm that will be used to borrow from the external data". Then suggest adding one additional sentence: "Examples of useful adaptive methods for cautious historical data borrowing include power priors, commensurate priors, robust mixture priors, elastic priors, and variants of these techniques that incorporate propensity matching to provide additional protection against unwarranted borrowing (e.g., due to unanticipated bias in the historical data)."
EFPIA	771	772	5.35	Any prior used should be evaluated from a clinical perspective for plausibility in the contect of the chosen external data. Simulations should be chosen in alignment with this (e.g. if the 99% credible interval corresponds to some response rate, then not much larger response rates would need to be studied). Moreover the prior should be discussed from an endpoint perspective. That is, the scale of the endpoint should be taken into account when considering the prior (e.g. log-odds are realistically between -4 and 4 for usual rersponse rates and respectivley for other endpoints or event rates are meaningfully only in units of years possibly - this depends on the context.)	Please include additional considerations and aspects to faciltiate transparent documentation of a prior which are non-technical.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	771	772	5.35	This text could be misinterpreted as uniformly suggesting a static approach to borrowing. It is often a poor decision to pre-specify a fixed, static “amount of borrowing from the external data”.	The recommendation should be that the sponsors pre-specify “the exact and quantitative approach to borrowing, including whether borrowing is static or dynamic, the structure of the inferential model, and all prior probability distributions” or something similar. It would also be useful to explicitly state that the choice of prior distributions in hierarchical models, e.g., used for dynamic borrowing, should generally be supported by simulations evaluating operating characteristics.
ACRO	772	772	5.3	Challenge: The discussion of borrowing from external data and trial success is unclear.	Recommendation:
ACRO	772	772	5.3		ACRO recommends highlighting the main types/methods/examples of acceptable borrowing of external information in the context of adaptive designs using Bayesian Borrowing method. It would be helpful to add examples for different extents/amounts of borrowing external data such as hybrid control arms and/or full external comparator/control arms. We suggest the following additional text at the end of line 772:
ACRO	772	772	5.3		“The amount of borrowed external data could range from complementary prior data to current trial control arm (i.e. hybrid control arms) to full external comparator/control arms.”
MRC Clinical Trials Unit	773	774	5.3	Requiring "control of the chances of false positive conclusions" conflicts with the previous paragraph which advocates a clinically defensible prior. It is typically not possible to both express clinical beliefs and control type 1 errors.	
EFPIA	774	774	5.35	The need to control type-1 error using a Bayesian Framework may not be an appropriate approach to risk management in the situation, where borrowing shall be implemented. Using dynamic borrowing, the type-1 error could be controlled at a level as for a frequentist framework for most meaningful scenarios, while it is increased for scenarios, which are considered a-prior less likely.	Remove: ", including control of the chances of false positive conclusions" or replace with an appropriate formulation in the Bayesian Framework considered.
EFPIA	774	774	5.35	Here we should also clarify that the Bayesian approach hinges on the trial design. Hence, the statement "control of the chances of false positive conclusions" should be refined.	Add here the reference to the trial design like "control of the chances of false positive conclusions for the trials design".
EFSPI/PSI Regulatory ESIG	774	774	5.3	"control of the chances of false positive conclusions " This is an imprecise formulation. Does it refer to Type I error? Or something else? If type I error that immediately exclused many Bayesian methods.	Make the term "control of the chances of false positive conclusions " more precise.
Invents consortium - EU Horizon project	774	774	5.3		It is unclear if the “chance” here refers to a Bayesian probability (averaging over the prior) or an error under the null hypothesis. This should be made clear (here and in other places in this document).
EFPIA	775	775	5.35	It is not clear what is "the balance between the prior and trial data"	Suggest to amend to: "the balance on utilized amount of information between the prior and trial data"

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	777	779	5.35	Conflict is not a well defined concept. For example, if the true placebo group effect is the 25 or 75 percentile of the prior, it is not in conflict with the prior. However, others might view these scenarios in conflict and regardless, error control should be investigated. Also, if adopting robust priors, they reduce the possibility of conflict (or some would argue virtually eliminate). Would suggest the concept is defined or another way to express the point is developed.	Define the concept or provide another way to express the point.
IQVIA	777	777	5.3	Highlighting the main types or models of acceptable borrowing of external data in the context of adaptive designs would provide reassurance of acceptability of data borrowing as potential solutions when a representative sample size has not been achieved or is not achievable, especially in rare diseases.	At the end of line 772, suggest adding examples for different extents or amounts of borrowing external data; e.g.: hybrid control arms and/or full external comparator/control arms. Suggested text to be added: "The amount of borrowed external data could range from complementary prior data to current trial control arm (i.e.: hybrid control arms) to full external comparator/control arms."
PhaseV Trials, Inc.	784	785	5.3	I was disappointed that this document stopped short of discussing designs where Bayesian tools are being used for Bayesian goals -- say, to quantify the posterior probability that a treatment effect exceeds some clincally relevant threshold. Such designs are often more readily interpretable by clinicians that those based on traditional p-values, and provide full posterior (and predictive) inference on all model quantities of interest. Such designs need not involve informative priors, nor borrowing from historical data; that is, they are no more "subjective" than traditional frequentist analyses. Indeed, analyses of this type are currently being encouraged by FDA in their Bayesian Statistical Analysis (BSA) Demonstration Project (https://www.fda.gov/about-fda/cder-center-clinical-trial-innovation-c3ti/bayesian-statistical-analysis-bsa-demonstration-project). Studies eligible for this program must use a fairly simple, Phase 3 design (e.g, no external data borrowing), and may use Bayesian methods for a pre-specified primary analysis, for supplemental analysis of the primary endpoint in the overall study population and/or in relevant subgroups, or for trial monitoring.	Suggest adding a final paragraph to this section describing the use of Bayeisan methods for Bayesian goals, in settings where frequentist methods may also be appropriate, but where posterior or predictive summaries are most naturally used to summarize trial results. It can be emphasized that such approaches will typically use non- or minimally-informative prior distributions, since the point here is not to borrow from historical information, but rather to take advantage of the flexbility and generality of the Bayesian paradigm. Careful justification for the use of Bayesian methods, as well as corresponding sensitivity or "tipping point" analyses, can certainly still be required.
EFPIA	784	785	5.35	If the planned trial with borrowing adheres to the "Key principles" outlined in Section 3 and the additional considerations outlined in Section 5.3, then it is not clear why the trial also needs to be evaluated with no borrowing. A trial without borrowing may be underpowered in such a situation and not be adequate. It is also unclear how such an approach would be implemented in cases with a virtual control or borrowing of an entire control arm.	Remove "It is also important to evaluate the current trial data with no borrowing".
EFPIA	786	804	5.4	For a time-to-event endpoint, the treatment may have a delayed effect. Thus, the proportional hazards assumption will not hold and the estimate of treatment effect from an interim analysis may underestimate the overall treatment effect. Adaptation solely based on the estimate of treatment effect from the interim analysis may have issues. Other information should be incorporated in the adaptation decision making.	Consider adding the following sentence in line 804: "For time-to-event endpoints, adaptation decisions should not rely solely on an interim treatment effect estimate given the possibility of delayed treatment effect. Additional information should be incorporated into adaptation decision-making."
c4c-S	786	824	7.8	Time-to-Event Designs	Correct identification of issues with independence and Type I error. Recommend adding methodological reference (e.g., Bauer & Köhne combination test) for appropriate statistical handling. Clarify expectations for censoring handling at adaptation points.
c4c-S	786	824	5.4	No comments	
EFSPI/PSI Regulatory ESIG	788	789	5.4	The statement "In such time-to-event trials, the statistical power of the trial depends on the number of events rather than the number of participants" is true for logrank and other HR-based tests but not necessarily if one is interested in other approaches, say RMST.	Perhaps consider slightly re-phrasing the sentence to address the concern.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFSPI/PSI Regulatory ESIG	789	789	5.4	As this section is rather about the time to a first event and individuals may have more than one event of the outcome of interest, consider using "number of participants with an event".	Please consider replacing "number of events" by "number of participants with an event" throughout this section?
MRC Clinical Trials Unit	789	795	5.4	Some multi-arm trials validly use control arm events rather than total events to time analyses (doi:10.1002/sim.1430).	Change "all participants are followed until a certain number of events have occurred" to "all participants are followed until a certain number of events (overall or in the control arm) have occurred".
nQuery	793	800	5.4	Lines 793-800 reference that in sample size re-estimation for survival analysis that an increased number of required events can be reached by either increasing follow-up of existing subjects or increasing sample size (or mix thereof). However, the choice between these strategies can induce a bias against or in favour treatment in the presence of a time varying effect. This was demonstrated in Freidlin & Korn (2017) (https://doi.org/10.1177/1740774517724746) for various scenarios via simulation. Given the increasing prevalence of complex non-proportional hazard patterns (e.g. delayed effects seen commonly in immunotherapy) in oncology trials, this means the choice of sample size re-estimation strategy could have a substantial effect on bias (e.g. in presense of delayed effect, there would be a strong sponsor incentive to favour increasing follow-up strategy) and therefore is worthy of specific note in ICH E20	Emphasis on pre-specification of SSR strategy and its potential impact Extra caution/explicity do no recommend SSR if reasonable prediction of complex survival patterns
EFPIA	805	824	5.42	This paragaraph not only refers to "Adaptive Designs in Time-to-Event Settings" but also to longitudinal settings. Therefore, it would be more appropriate to create a new section for this paragraph with a title such as "Adaptive Desings with a Potential of Depencence of Data between and after Adaptations"	
EFPIA	808	824	5.42	In discussing the considerations when a patient may contribute information both before and after an interim analysis, the draft Guideline should explicitly mention the potential value of simulation of this data structure and timing to quantify the effects, if any, on operating characteristics including type I error control. In many cases, the quantitative effects—while real—are of an insufficient magnitude to constitute a meaningful threat to trial validity.	Suggested text for a new paragraph, to be added after the paragraph that ends on line 824: "Alternatively, there may be value in simulating the proposed trial and associated data structure, including the data from participants who contribute information both before and after an interim analysis, to quantify the effects, if any, on operating characteristics including type I error control. In many cases, the quantitative effects—while real—are of an insufficient magnitude to constitute a meaningful threat to trial validity."
EFSPI/PSI Regulatory ESIG	821	824	5.4	This is a separate and important considerations that perhaps is somewhat hidden (or misplaced) at the end of this section.	Please consider putting this statement earlier, perhaps mentioning so-called "pipeline" participants?
Ferring Pharmaceuticals	821	824		Hampson and Jennison (2013) J. R. Statist. Soc.: 75, Part 1, pp. 3–54, show that for group sequential designs (allowing for stopping for futility or efficacy only), incorporating a short-term endpoint or intermediate outcome measurements in a repeated measurements setting, the independent increments assumption does hold. If the authors of the ICH E20 guidance agree with this conclusion, it would be helpful to clarify that the mentioned conceptual problems do not arise in simple group sequential designs. Otherwise, confusion might arise as to the appropriateness of using standard group sequential designs in this setting.	
c4c-S	825	853	8	Adaptive Designs in Exploratory Trials	suggest adding warning that exploratory adaptation decisions should not be retroactively justified in confirmatory settings. Recommend clearer distinction between exploratory flexibility and confirmatory rigor.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
c4c-S	825	853	5.5	No comments	
EFPIA	826	834	5.51	For exploratory trials, the statement “it is critical that principles in section 3 are followed” could result in more Phase 2 trials being designed as a “mini Phase 3”. Objectives, uncertainties and hence designs for exploratory trials may differ severely from confirmatory trials. In particular, the use of augmented controls and Bayesian adaptive designs are frequently considered in early Phase development. The statement may introduce misinterpretations complicating discussions within industry and between industry and regulators as any RCT could potentially support benefit-risk assessment.	Suggestion is to just add this sentence at the end of the first paragraph and further clarify that this statement only is of relevance for trials, which should serve as confirmatory evidence: "This guideline focuses on the use of adaptive designs in confirmatory clinical trials. Adaptive designs may also be used in exploratory trials early in drug development that are intended to obtain information on a wide range of aspects of treatment use (e.g., choices of dose, regimen, population, endpoints). Trials at this stage of the development program may include a larger number of adaptations to generate information that support important decisions about subsequent development phases. The principles in this guideline are also relevant in these settings to ensure the reliability and interpretability of the results and subsequent decision-making based on such trials. If an exploratory trial is intended to also confirm efficacy, it is critical that the principles in Section 3 are followed."
EFSPI	826	834	5.5	For exploratory trials, the statement “it is critical that principles in section 3 are followed” could result in more Phase 2 trials being designed as a “mini Phase 3”. Objectives, uncertainties and hence designs for exploratory trials may differ severely from confirmatory trials. In particular, the use of augmented controls and Bayesian adaptive designs are frequently considered in early Phase development. The statement may introduce misinterpretations complicating discussions within industry and between industry and regulators as any RCT could potentially support benefit-risk assessment.	Suggestion is to just add this sentence at the end of the first paragraph and further clarify that this statement only is of relevance for trials, which should serve as confirmatory evidence: "This guideline focuses on the use of adaptive designs in confirmatory clinical trials. Adaptive designs may also be used in exploratory trials early in drug development that are intended to obtain information on a wide range of aspects of treatment use (e.g., choices of dose, regimen, population, endpoints). Trials at this stage of the development program may include a larger number of adaptations to generate information that support important decisions about subsequent development phases. The principles in this guideline are also relevant in these settings to ensure the reliability and interpretability of the results and subsequent decision-making based on such trials. If an exploratory trial is intended to also confirm efficacy, it is critical that the principles in Section 3 are followed."
EUCROF - EU CRO	826	828	5.5	"This guideline focuses on the use of adaptive designs in confirmatory clinical trials. If a trial may be intended to confirm efficacy and support benefit-risk assessment, it is critical that the principles in Section 3 are followed."	**Proposed Change:** Please add sentence after "… the principles in Section 3 are followed": "After careful consideration of introducing possible bias, a **limited number of adaptations** can be accepted in a confirmatory trial"
Invents consortium - EU Horizon project	826	828	5.5	This guideline focuses on the use of adaptive designs in confirmatory clinical trials. If a trial may be intended to confirm efficacy and support benefit-risk assessment, it is critical that the principles in Section 3 are followed.	A limited number of adaptations can be accepted in a confirmatory trial.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
Breakthrough T1D	828	832	5.5	In cell therapies, there is a need to consider variety of factors in the early stage to establish the appropriate dose and sites of implantation. This may also require the consideration of ancillary factors such as the scaffold/matrix (acceptability, grafting efficiency, type of tubing/material etc.). However, as mentioned earlier cell therapy trials are currently small in size, single-arm and unblinded. It would be most helpful if the EMA could request the ICH to consider tailored adaptive designs in early exploratory trials for cell therapies balancing the need for establishing baseline conditions vs. the smaller set of participants in clinical trials.	
IQVIA	830	830	5.5	Reinforcing the message that simulations are also applicable and acceptable in exploratory trials is important, not just for development productivity, but more so for the safety and benefit of participants. Simulations can be valuable in early phases as they help exclude trials earlier that are likely to fail, reducing the number of participants exposed unnecessarily to risks of an ineffective investigational treatment and providing participants the opportunity to explore more promising	After the period (end of sentence) within line 832, suggest adding: "Simulations can help reduce the number of adaptations required in such exploratory trials."
ACRO	832	832	5.5	Challenge: ACRO notes that simulation studies, as described in section 5.2, are also applicable. As simulation studies may help to refine and optimize exploratory trials, it would be helpful to add this into this section.	Recommendation: We suggest this additional sentence for the final guideline: "Simulations can help limit/reduce the number of adaptations required in such early/exploratory trials."
EFPIA	832	834	5.51	Section 5.5 refers to exploratory trials. It is mentioned that the guidance principles are also relevant in that setting. Still, some flexibility is permitted (no need for strict specification of adaptation rules, sponsor's role in interim decision making can be different,...). However, the expectation in terms of controlling the chance of erroneous conclusions in an exploratory setting is not addressed. Some flexibility also permitted for that key principle.	Clarify the expecations in terms of controlling the chances of erroneous conclusions in the context of exploratory trials
EFPIA	833	834	5.51	The guidance states that "The principles in this guideline are also relevant in these settings to ensure reliability and interpretability of the results and subsequent decision-making based on such trials." However, the principles mentioned in this guidance are focused on controlling the statistical properties such as the type I error rate and maintaining trial integritiy in the trial. Exploratory trials such as early phase dose finding trials are non-randomized and follows a deterministic procedure for dose escalation stage and even when patients are randomized to two or more doses after the dose escalation stage, the sample size is selected through clinical considerations and controlling the type I error rate is not the main focus here. Therefore, the principles of adaptive designs are quite different for such exploratory trials. The guidance document, when it speaks about exploratory trials needs to be specific about this point when mentioning the principles of adaptive designs in such exploratory trials.	Suggest having a mention in the document that maintaining the type I error rate in exploratory trials such as the dose finding trials is not the focus. We need to have enough patients to be able to assess the additional data needed to identify the optimal dose.
c4c-S	854	880	5.6	No comments	
c4c-S	855	880	5.6	The text is difficult to read.	I would recommend to add numbers to the examples of challenges, i.e. 1-measures should..., 2-informed consent forms, etc.
EFPIA	856	862	5.61	For many adaptations, it would be unethical to withhold their implementation from trial participants. These ethical considerations must be balanced against the need to minimize the dissemination of sensitive interim information information.	Section 5.6 should include a discussion on the operational aspects of maintaining trial integrity under consideration of informing trial participants about key changes to the study.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	858	862	5.61	Sponsors of clinical trials have an obligation to ensure that both currrent and prospective participants provide fully informed consent. The status of a trial following an interim analysis constitutes relevant information, and therefore it may be argued that the results of such analyses should be disclosed, at least in general terms. For instance,if a trial has "survived" a futility interim, this implies that some preliminary evidence of efficacy has been observed. Conversely, the continuation of a trial after an efficacy interim suggests that the observed efficacy has not yet reached a predefined threshold. Guidance would be valuable on two aspects: (1) when and how the Participant/Patient Information Sheet (PIS) should be amended - and re-consent obtianed from current participants - and (2) the extent of information that may appropriately be shared with participants following a non-terminal interim analysis.	Suggest to include additional text discussing the impications of conducting an interim analysis on the informed consent process.
EFPIA	858	859	5.61	The description of possibly complex adaptive elements in an informed consent document may be confusing and even misleading to prospective trial participants, so their inclusion should be considered on a case-by-case basis by the appropriate ethics committees or equivalent. For the individual prospective participant, while it is critically important that they be informed regarding the goals of the trial and current state of knowledge, what may happen later in the trial may be largely irrelevant to their own benefit-risk evaluation. For example, for a prospective participant considering enrollment in the first stage of an adaptive trial with population enrichment, it is important that they are informed that it is unknown whether there will be benefit but it may not be useful to know that the inclusion/exclusion criteria may be changed, possibly years later after their involvement is completed. In some cases, it may be appropriate or necessary to modify the informed consent document after an adaptation, e.g., if an active arm is dropped from a multi-arm trial, but the possibility of that adaptation may not be information that is useful to prospective participants.	The text suggesting the including of complex--and possibly minimally relevant--adaptive elements in the informed consent document should be removed. It may be useful to state that the inclusion of methodological detail in the informed concent document should be determined on a case-by-case basis by the appropriate ethics committees and other personnel, with the goal of ensuring prospective participants have all information required to assess their own risks in participating in the trial.
EFSPI/PSI Regulatory ESIG	864	866	5.6	It is stated that "[c]linical trials with an adaptive design typically use an interactive voice or web randomization system to manage randomization and assignment of participants to treatment arms." We think it would be important to mention that these systems are in fact used in almost all global clinical trials, not only those with an adaptive design. We however acknowledge the fact that it is even more important to use an IRT system for trials with adaptive design features due to their increased complexity, not only regarding randomization but also drug supply and other critical trial elements. In addition, we would advise to use the term Interactive Response Technology (IRT) systems or IxRS (the x standing for either voice or web) instead of IVRS/IWRS, as this is more common terminology nowadays.	We recommend to use the term "Interactive Response Technology (IRT)".
EFPIA	868	879	5.61	mention that changes in the treatment arms or randomization ratio should be done with "minimum sponsor involvement", which contradicts the statement in lines	Suggest revisiting this statement and modifying accordingly
EFSPI/PSI Regulatory ESIG	868	870	5.6	It is mentioned here that changes in the treatment arms or randomization ratio should be done with "minimum sponsor involvement", which contradicts the statement in lines 549-550 which states that these changes should be done "without sponsor involvement". While the statement in 549-550 goes against the ICH E6(R3) guidance that the sponsor should conduct oversight of trial-related activities, minimum sponsor involvement seems to be the only feasible approach when the sponsor is also to fulfill oversight requirements. The guideline should have a consistent position on this topic, and also mention the oversight requirements laid down in ICH E6(R3).	We would appreciate consistent recommendations - within and across guidelines.

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFPIA	870	880	5.61	Adaptations in clinical trials inherently create unpredictable demands on drug supply and can potentially lead to delays. This represents a great opportunity to emphasise the role of operations in adaptive designs. For instance, a stronger push should be advocated on evaluating the operational characteristics—such as the adaptation's impact on drug supply, randomisation, data quality and MRCT considerations—when weighing the pros and cons of adaptive designs versus alternative designs. Furthermore, the concept of a "design for adaptive operations", should be promoted, ensuring that operational aspects are planned and aligned with similar level of statistically rigor present throughout this guidance document.	Propose to add the following paragraph to this section: "Adaptations in clinical trials inherently introduce unpredictable demands on drug supply, potentially causing delays that adversely impact trial timelines. This challenge presents a significant opportunity to underscore the crucial role of operational planning in adaptive design strategies. An enhanced focus should be placed on thoroughly assessing the operational characteristics—specifically examining the impacts of adaptations on drug supply, randomisation processes, data quality, and considerations for Multi-Regional Clinical Trials (MRCT), Decentralised Clinical Trials (DCT) and Pragmatic Clinical Trials (PCT). When evaluating the benefits and drawbacks of adaptive designs in contrast to traditional approaches, these operational aspects should be considered. Moreover, the introduction of a 'design for adaptive operations' should be championed, ensuring that operational planning is as rigorously conceived as the statistical methods detailed in this guidance document. Doing so will help align operational strategies with statistical rigor to optimise trial efficiency and mitigate risks associated with supply fluctuations and other operational challenges."
EFSPI/PSI Regulatory ESIG	879	880	5.6	The term "formal interim database lock" is used. It would be important to specify how a "formal interim database lock" is defined and how it is distinguished from an "informal interim database lock". ICH E9 specifically defines an "interim analysis" as any analysis intended to compare treatment arms with respect to efficacy or safety at any time prior to formal completion of a trial" and does not distinguish "formal" or "not-formal" locks. Further clarification for the use of the term "formal" therefore is requested.	We would appreciate a definition of "formal" and "informal" data base locks.
EFPIA	881	952	6	Section 6 provides guidance on how adaptive designs should be documented prior to the conduct and which information needs to be included in the marketing application. However, no guidance is given on whether or how adaptations are to be documented while the trial in ongoing.	Add section on documention requirements during the conduct of a trial with an adaptive design.
c4c-S	881	928	6.1	No comments	
ACRO	882	928	6.1	Challenge: The section is thorough but risks creating redundant documentation across regulatory regions. There should be reassurance that adaptive details can be incorporated into existing core trial documents, not as a separate new deliverable.	Recommendation: We recommend the addition of new sentence in line 922: "Details regarding adaptive elements should be included within the appropriate existing study documentation - such as the protocol, statistical analysis plan (SAP), or data management plan (DMP) - rather than requiring a new standalone 'Adaptive Design Plan.'"
ACRO	882	928	6.1	Challenge: The section could better align with ongoing ICH digital harmonization efforts (M11 protocol template, M13 data submission standards).	Recommendation: We suggest the addition of new sentence in line 920: "Documentation expectations should be harmonized with ICH M11 and regional data standards to support structured, digital transparency and reduce duplication across submissions."
EFPIA	885	885	6.11	The text states: "In addition to the information typically include in a clinical trial protocol or in other documents, where suitable documentation should include the following". It is recommended that further guidance provided on where this information might be included.	Recommendation to specify clearly where this information is expected by regulators (e.g. SAP, interim SAP, Simulation report).

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
EFSPI/PSI Regulatory ESIG	893	896		Data cleaning for interim analyses is a major operational consideration. Data cleaning requires cross-functional input and co-ordination of resources and timelines so should be discussed as early in the study as possible to allow appropriate discussions take place, with time to implement decisions made. This merits a separate paragraph to draw attention to this critical to quality process.	Suggest to move data cleaning details to a separate paragraph, and emphasising the importance of having this discussion early to allow sufficient time to resolve budget and resource constraints. Add a note that the accumulation of data should be carefully monitored to ensure that patients included in the interim analysis have sufficient data. For example, the necessary number of patients might have reached the timepoint to trigger the interim anaysis but a larger than expected proportion might have key data missing or collected outside of the protocol specified window, impacting the reliability of the interim analysis results. Early identification of this potential risk should be built into data monitoring plans.
c4c-S	929	945	9	Post-Trial Documentation (Marketing Application)	suggest requiring side-by-side display of planned vs. implemented adaptations with rationale for deviations.
c4c-S	929	952	6.2	In agreement with Maria (c4c-S)'s recommendations	cf Maria (c4c-S)'s recommendations
Teva Pharmaceuticals	929	952	6.2	• Section 6.2 lists down the documentation that needs to be included in a marketing application following the completion of a confirmatory trial with an adaptive design. Clarification is needed where in the Marketing Authorisation Application these documents should be located. The current ICH E3 guideline does not address all the documentation specified in this section.	
EFPIA	933	933	6.21	It is unclear where in the marketing application this information is to be included.	Recommendation to specify clearly where this information is expected by regulators (e.g. SAP, interim SAP, Simulation report, study report).
EFPIA	935	941	6.22	While Section 6.2 mentions reporting "whether anticipated adaptation rules were followed," it could be strengthened by explicitly requiring a detailed justification for any deviations from the pre-specified rule.	Modify point 2 in Section 6.2 to read: "...the adaptation decisions that were made, whether anticipated adaptation rules were followed, and a detailed rationale for any deviations from those rules." This adds an explicit requirement for justification, which is critical for regulatory review.
EFSPI/PSI Regulatory ESIG	941	941	6.5	The term "date of sponsor unblinding" is mentioned, but depending on the needs of the trial, there may be multiple dates of sponsor unblinding varying by function. If by sponsor unblinding the term "database unblinding" is meant, which can be tied to an actual distinct date, then the latter term should be used instead.	Please clarify what is meant by "date of sponsor unblinding"
c4c-S	946	948	9	Integrity Compliance Reporting	Consider recommending independent audit certification of data access logs before submission.
EFPIA	949	950	6.25	The word "record" used twice could be mis-understood as audio report or full verbatim of the discussion, which may not be feasible.	"Records of deliberations by the IDMC (e.g., all closed and open IDMC meeting minutes), including ~~records~~ minutes of ~~discussions~~ related to any adaptation decisions."

Name of organisation or individual	Line from	Line to	Section number	Comment and rationale	Proposed changes / recommendation
c4c-S	949	952	9	IDMC Records and Adaptive Reporting	Important transparency section. Suggest standardization of IDMC documentation (e.g., redacted minutes for regulators, protected minutes for archive). Also recommend requiring data provenance statement confirming interim-to-final data linkage accuracy.
EFSPI/PSI Regulatory ESIG	951	951	6.2	"Reporting of results that appropriately account for the adaptive design (e.g., appropriately adjusted estimates, confidence intervals, and p-values)." That is of course desirable. However, methods for adjustment may not exist. Does that then mean a design is not eligible?	Clarify whether this is mandatory, and what to do if methods are not available.
Dr.Viviana Mascilongo	953	1000		Missing acronym legend and bibliography	