CLINICAL INVESTIGATION

INTERVIEW

Clin. Invest. (2012) 2(3), 235–240

Perspective on adaptive designs: 4 years European Medicines Agency reflection paper, 1 year draft US FDA guidance – where are we now?

"...the importance of speeding up (or optimizing) drug development is well understood. The challenge, however, is to arrive there without lowering established standards for drug licensing..."

Adaptive clinical trials attract great attention from academia, industry and regulatory authorities. Both the European Medicines Agency and the US FDA have clarified their positions in recently issued (final or draft) guidance documents. With this background, current trends and issues were analyzed in a panel discussion at the International Society for Biopharmaceutical Statistics (ISBS) meeting in March 2011. In this article, members of the panel summarize their thoughts based on this discussion.

Keywords: adaptive design • group-sequential trial • regulatory guideline

Brenda Gaydos[1], Armin Koch[2], Frank Miller[3], Martin Posch[4], Marc Vandemeulebroecke*[5] & Sue-Jane Wang[6]

[1]Eli Lilly & Company, Indianapolis, IN, USA
[2]Medizinische Hochschule Hannover, Germany
[3]AstraZeneca R&D, Södertälje, Sweden
[4]European Medicines Agency, London, UK
[5]Novartis Pharma AG, Basel, Switzerland
[6]Office of Biostatistics, Office of Translational Sciences, Center for Drug Evaluation & Research, US FDA, Silver Spring, MD, USA
*Author for correspondence:
E-mail: marc.vandemeulebroecke@novartis.com

Adaptive clinical trial designs continue to attract great attention from academia, industry and regulatory authorities. Methodological research has been prolific and a number of practical applications have followed. From the regulators' side, two position papers have recently been issued: the European Medicines Agency (EMA) Reflection Paper on 'Methodological issues in confirmatory clinical trials with flexible design and analysis plan' [1] and the draft US FDA guidance on 'Adaptive Design Clinical Trials for Drugs and Biologics' [2].

Both documents have sparked fruitful discussions after their appearance; for example, most recently in a special issue of the *Journal of Biopharmaceutical Statistics* [3] focusing mainly on the recently issued draft FDA guidance. The International Society for Biopharmaceutical Statistics (ISBS) meeting in March 2011, jointly organized by the EMA and the German Region of the International Biometric Society (IBS), provided the opportunity to discuss the latest trends and issues in an open forum. What burning questions still persist – and can we answer them now? Where are there still gaps or opportunities? A panel discussion was held with key individuals from the US and European authorities and other statisticians actively involved in the field of adaptive clinical trials were invited to comment on these and other questions. In this article, members of the panel summarize their thoughts based on this discussion.

Q Under what circumstances does the regulator consider an adaptive trial 'adequate and well-controlled' (A&WC) in the sense of the draft FDA guidance and/or 'pivotal'?

Sue-Jane Wang: By legal statue in the USA, a clinical trial, whether adaptive or not, is considered A&WC if it satisfies those criteria defined in Title 21 of the Code of Federal Regulations under Section 314.126 'Adequate and well-controlled studies.' For a confirmatory adaptive trial, a strong control of studywise type I error rate should be a minimum criterion. In addition, the study should contain the three components

ISSN 2041-6792 10.4155/CLI.12.5

described in Section III.A of the draft FDA guidance, namely, prospectively planned opportunity for adaptation, prospectively planned time point for adaptation and the interim analyses can be performed blindedly or unblindedly, and with or without formal statistical hypothesis testing [2]. There should be consistency in the descriptions among study design, study protocol, statistical analysis plan and study reports. The criteria for consideration as A&WC include, but are not limited to, study objectives, analysis methods, methods used for patient selection, patient assignment, study bias, comparative evidence, study conduct, study end points and interpretability of analysis results.

Armin Koch: The European Reflection Paper uses the term 'confirmatory trial', which has been defined in Section 2.1.2 of the internationally agreed guideline 'Statistical Principles for Clinical Trials' [4]. This definition covers the aforementioned principles and gives the same emphasis to consistency between the preplanned primary objective(s), the primary hypotheses, the conduct and the analysis of a clinical trial. Preplanning is the key to distinguishing between exploration and confirmation and in consequence, adaptation in confirmatory clinical trials is only possible if at least the type of the adaptation can be clearly preplanned (and the need to do so, should be justified). It is obvious that the option to adapt the design challenges the traditional concept of phases of drug development with exploration and confirmation, but this concept is still helpful as a yardstick with which the overall evidence provided for licensing should be measured.

Brenda Gaydos: SJ Wang already referred to the definition of A&WC in 21CFR314.126. The key point is that this definition applies equally to all designs, including adaptive designs. Also, any adaptive dose-finding study that meets the criteria for A&WC should be able to serve as pivotal evidence of effectiveness, similar to any other design that meets those criteria. An example is a two-staged inferentially seamless Phase II/III treatment selection design. This general class of designs allows for learning about the dose response in stage 1 and confirming the targeted doses selected for continuation in stage 2 by combining information across stages. There are statistical methods (e.g., combination function approaches) that strongly control the type I error rate and allow full flexibility in how doses are selected at the end of stage 1.

Marc Vandemeulebroecke: I can see two situations where an adaptive trial might be more easily accepted as 'pivotal.' First, in special situations such as orphan diseases or when ethical constraints dictate some level of flexibility in a trial. The EMA Reflection Paper uses the wording 'difficult experimental conditions.' Second, in the context of the 'two pivotal trials paradigm', the sponsor might seek a discussion about the scenario of conducting one adaptive Phase II/III trial for final-dose selection and another conventional confirmatory trial.

Martin Posch: The EMA Reflection Paper lists a number of specific requirements for an adaptive design to be considered as confirmatory, such as the prespecification of the interim analysis and the type of adaptations, and the control of the information flow. Furthermore, it must be ruled out that the interim analysis introduces bias by implicit adaptations; for example, through a change in the characteristics of the patients recruited after the interim analysis, or a change in the assessment of the end point. Additionally, the requirements for classical fixed sample clinical trials apply equally to adaptive trials.

Q Assume an adaptive Phase II trial with, say, seven dose groups and very convincing results for some of the doses. How do we analyze this? Can such a trial be used for filing for approval?

S-JW: It seems to me the analysis will depend on the prespecified adaptation rules for dose selection and how the data are to be used in each of the stages as an adaptive dose-selection design. Although statistical methods are available, some with strong control of the type I error rate and some without, whether such a trial can be used for filing for approval would likely depend on the role this trial plays in the drug-development program. It might be difficult to articulate its confirmatory intent if still including many doses. In some cases, it might be possible that the results of the adaptive trial provide supportive evidence if clear evidences in confirmatory trials have been shown.

MP: The level of evidence provided by such a trial will depend on the specifics of the trial design. For very complex adaptive designs (e.g., response–adaptive designs), a rigorous validation of the statistical operating characteristics can be challenging and the potential sources of bias are more difficult to control. In contrast, if a limited number of doses are under investigation and adaptations are limited to a small number of predefined interim analyses, the requirements for confirmatory trials can be more easily satisfied.

AK: The European Reflection Paper is restricted to the discussion about adaptation in confirmatory clinical trials. While an early Phase II trial with seven doses can provide valuable information that may be urgently needed in the overall development program, this is not what has been foreseen as the type of dose finding within a confirmatory clinical trial. Dose selection in a confirmatory setting refers more to situations with a narrow dose range, where earlier stages of drug development suggest, say, two or three promising doses, and more information about tolerability or safety is required to make a final

decision about which dose to license. Dropping doses that are inferior in this aspect should be possible at the earliest time point. Ideally, these combined Phase II/III trials are performed in situations where the same end point is investigated in Phase II and III trials and conducting another Phase II study would not have further advantages beyond eventually making a better guess about the optimal dose from a safety perspective.

Frank Miller: In a Phase II dose-finding study with several, say, seven, doses the overall picture needs to be convincing; for example, if doses two and four are just statistically significantly different to placebo (after multiplicity adjustment) and the observed result for doses one, three, five, six and seven is similar to placebo, this would raise concerns. Therefore, a test analyzing the overall results to show an upward trend across doses seems reasonable, and is furthermore recommended in the International Conference on Harmonization E4 guideline [5]. For nonadaptive dose-finding studies, different tests have been described to show an upward trend across doses and some of them do not rely on a specific assumed dose–response shape. These tests can be applied in an adaptive dose-finding study as well, but an adjustment is needed.

BG: Approval depends on the full data package and typically a minimum of two A&WC trials are needed. There are several ways to analyze such a design (e.g., trend test, pairwise comparisons, fit a model and estimate the response at each dose from the model). One approach is to design the trial and analysis with the necessary precision to estimate the dose–response and select doses with confidence for Phase III (or to stop development). It is not likely that a sponsor would power a seven-arm study and strongly control the type I error rate for all of the hypotheses (e.g., pairwise comparisons to control), but if so then it should be possible that a design with seven doses would meet the criteria for A&WC. Although, given the number of doses, it may be questioned whether enough was known about the dose–response prior to starting the confirmatory trial.

Q Both guidelines focus on confirmatory drug development, but mention early development as a strong arena of opportunity for adaptive trials. What are our current experiences in confirmatory drug development? And in early development?

BG: Most experience in confirmatory development is with traditional group sequential designs and sample size re-estimation (both blinded and unblinded) and then next in two-stage treatment selection designs. There is increasing use of Bayesian methods for adaptive decision making within the frequentist framework (i.e., flexible adaptive designs that enable strong control of the type I error rate, independently of how the adaptive decision is made). Since the FDA draft guidance has come out, I have seen an increased interest in blinded sample-size re-estimation approaches in Phase III and adaptive dose-finding approaches in Phase II, both areas where the guidance was encouraging their use. In early development, experience is increasing; for example, with dose-escalation, seamless proof-of-concept/Phase II and adaptive dose-finding Phase II designs. Common approaches include dropping or adding treatment arms and response adaptive randomization to optimize identification of target doses. Model-based approaches are increasingly being applied (instead of multiple testing procedures) for final analysis and interim decision making.

FM: Still the most frequently applied adaptive design in confirmatory drug development is the group sequential design. Cautiously, other types of adaptive designs are applied in early development. Experience from these studies was usually very positive: better information was obtained compared with nonadaptive options and even the economic benefit could be quantified. Of course, areas for improvement were also identified from these studies. To mention only one example, we had an extremely well recruiting adaptive trial in our company. We have learned that next time we need to invest more effort in investigating and simulating different recruitment scenarios prior to the study. Lessons of this type are very beneficial.

Q What is the role of type I error control in exploratory trials, be they adaptive or not? Isn't power more important than the type I error in these trials?

S-JW: I will comment on adaptive exploratory trials. Control of type II error in earlier phase research should be critical as there is a need to select just one or a few promising treatment doses from among several more doses for confirmatory trial planning [6]. An exploratory adaptive trial can have wide flexibility and it may not be necessary to require a strong control of the studywise type I error rate. The probability of correct decision for planning confirmatory trials and achieving medical product approval or the probability of correct selection of design aspects may be a better metric to evaluate exploratory adaptive trials because of large uncertainties [6]. With diligent exploration, the wide flexibility in exploratory adaptive trials may lead to further development or may support an early decision to terminate the drug development program.

MV: Type I error control is less important in exploratory trials, but it can serve as a 'sanity device' to protect against over-optimism. Also, the exploratory setting may be a good field to gain experience with novel

adaptive methods and their operating characteristics, such as the type I error.

MP: Even in exploratory phases of drug development, the assessment of error probabilities in statistical procedures is of importance for efficient decision making, and type I and II errors should be weighted appropriately. Although the probability of false-positive claims is controlled by the confirmatory part of the drug development program, controlling the type I error in earlier phases can prevent failures in confirmatory trials.

AK: In my somewhat relaxed understanding of the ICH E4 guidance [5], one of the main messages about dose finding may be phrased saying that nothing needs to be formally significant once establishing dose–response of a certain drug is intended; however, dose–response actually can prove efficacy of a drug and therefore investing more into a Phase II trial (so that, for example, a concept for the type I error exists to identify a dose that is superior to placebo) may lead to a trial that can be assessed as part of the confirmatory package. So exploration with control of a type I error can in some instances be efficient.

FM: From a company perspective, Phase I or II trials with some efficacy readout are used to justify investment into the next development phase. The type I error is the risk to agree to continued investment although the investigational compound is like a placebo. It is good practice for decision makers in companies to quantify and control this risk. The acceptable risk needs not be the traditional (one-sided) 2.5%; type I errors of 5–20% are common in early phase trials. But it is important to make these risks visible and to use proper statistical methodology (e.g., unbiased tests) to support internal investment decisions. I agree that power is very important, especially when bearing in mind that a whole series of trials in the early development program need to be positive before a compound arrives in Phase III.

BG: Exploratory trials (including Phase II trials not intended to be pivotal) should be designed to maximize the probability of making the right decision within your operational constraints (maximize the true-positive and true-negative rates of the decision). This can be thought of as increasing the power of getting the decision right. Clarity is needed on the risk level acceptable for a given sample size because trade-offs will be needed in the decision rule. For example, is it more acceptable to increase your likelihood of a false-positive or false-negative decision? These probabilities are inversely related. Power and type I error in the frequentist setting ($1-\beta$, α) are used in creating decision rules for a positive or negative decision, but they are not probabilities associated with making the right decision. They are conditional probabilities of a true-positive or false-positive decision, based on hypothesized true values of the unknown effect. Unless the trial is intended to be pivotal, strong control of the type I error rate is not needed. However, it is important to understand the power of making the right decision based on the decision rules implemented.

Q Possible design modifications range from well-recognized ones (e.g., sample size reviews) to more novel methods (e.g., dose-selection, adaptive randomization and so forth). At present, what are our experiences with such approaches – are there any strongly emerging lead themes or newly emerging focus areas? Are there topics that need further exploration or understanding, methodologically and/or operationally?

BG: Bayesian response–adaptive randomization for Phase II may be the most efficient approach to dose finding [7]. However, these trials are complex to design because there is a lot of flexibility in the selection of data sampling rules, allocation rules, early stopping rules, dose selection rules, models (dose–response and longitudinal) and prior definitions. These are also among the most difficult approaches to implement well. More exploration and experience is needed in this area.

FM: One example for emerging topics is heterogeneity between stages, which was intensively discussed some years ago with regard to confirmatory designs. In Phase II dose-finding trials, this issue seems especially important to consider: if some doses are used only in parts of a study where a different population was recruited and other doses are used during the whole study, comparison between doses might not be straightforward [8]. The PhRMA working group on adaptive dose ranging commented that 'the possibility of a shifting patient population in the trial is, perhaps, one of the greatest challenges concerning the utilization of adaptive designs in practice' [9]. When being aware of this potential risk, things can be done to minimize it both during the conduct of the trial and when analyzing it.

MV: One lead theme appears to be adaptive approaches to dose-ranging. B Gaydos mentioned Bayesian response–adaptive randomization, but there are other examples. Many methodological advances have been made in recent years, and an increasing number of case studies are being conducted. In this area, a great deal can be gained from an efficient approach and sponsors are less restrained than in Phase III. Another arena for development might be the integration of all logistical aspects of adaptive trials. We have IVRS, specialized software for planning and analysis, electronic data capture, and so forth, but the integration of all these aspects into one interacting network is still complicated in practice.

S-JW: From our regulatory review experience of confirmatory adaptive trials, it is worth noting that

the concept of a minimum clinically important effect is important in regulatory decision making for a test drug product, from the perspective of benefit versus risk. When the minimum clinically important effect can be prespecified at the design stage, it has important implications for sample size planning and sample size reviews; for example, should the hypothesis for testing rule out this minimum important effect and, if so, the statistical decision rule for asserting that the test drug has an effect at least as large as this minimum effect needs to be well defined at the design stage [10].

Q The draft FDA guidance dedicates more attention to modeling and simulation activities than the EMA Reflection Paper. Is there a need for revising the EMA Reflection Paper to encourage simulations and statistical modeling?

S-JW: The draft FDA guidance discusses the roles of clinical trial simulation in adaptive design planning and evaluation in Section VII.D. However, the focus is mainly on the simulation-related matters that may or may not include the modeling for assessing the adaptive design performance characteristics. It is worth noting that if modeling is considered, the purpose of modeling in exploratory trials is generally different from that in confirmatory trials. Some are used chiefly for planning, while others may be used for planning and/or data analysis [11]. In general, we are open to any proposals and we provide scientific review comments that bear regulatory context for drug development.

AK: I feel that in general there is no need for special encouragement from regulatory agencies for what could be done during the planning stage of a clinical trial or a development program. During planning, everything that could be done to achieve a better understanding of the situation at hand should be done, obviously. The same is true for encouragement regarding special statistical methodology or experimental design (like Bayesian methods, t-tests or adaptive designs). Choice of methods for certain situations must depend on a careful justification and previous experience, with certain methodology or designs having been used successfully in planning, analyzing and interpretation of trials.

MV: It may not be the regulator's role to encourage specific techniques or approaches. However, it is worth noting that the FDA maintains a dedicated pharmacometrics division; the EMA draws on expertise across Europe. Modeling and simulation can serve as a 'modern version of protocol planning' [12] or as a means for data analysis and knowledge integration, not only in the context of adaptive designs. The term 'model-based drug development' has been coined for this.

MP: The appropriateness of specific methods typically depends on the concrete setting in which they are applied so I am skeptical in giving a general encouragement. For example, simulations of adaptive clinical trials are a valuable tool for assessing the power of adaptive clinical trials and for comparing the efficiency of different types of adaptation strategies under different scenarios. However, they have limitations in the evaluation of the type I error rate and validity of estimates. For a thorough assessment that is not restricted to the global null hypothesis of no drug activity in any dose and end point involved, an (often very large) parameter space has to be covered by the simulations or least favorable configurations have to be identified, which may not be feasible for complex adaptation procedures.

Q What are examples of 'less well-understood' topics or methods as mentioned by the draft FDA guidance? Has our understanding recently progressed and, if so, how?

FM: There exist highly flexible dose-finding designs that probably are examples for designs being 'less well understood.' However, if limited adaptation is done and type I error control is convincingly proven, adaptive dose-finding designs or adaptive choice-of-population designs could be candidates to be classified as 'well understood' (e.g., designs with one interim analysis and prespecified selection criteria).

S-JW: The term 'less well understood' in the draft guidance means 'less regulatory experience in terms of drug approval based on adaptive trials designed to be considered as adequate and well controlled.' The draft guidance considers those adaptively designed clinical trials that use interim, unblinded accumulating data for adaptation evaluation to be 'less well understood.' It seems that depending on the treatment indications, for example, if demonstration of a dose–response profile including longitudinal responses is a critical component for a successful drug development, then a two-stage adaptive dose response and/or dose selection design for consideration as A&WC can be at great risks of the sponsor if early exploration has not been properly pursued.

MV: The dichotomy of 'well-understood' versus 'less well-understood' methods has already been commented on, for example, by Brannath *et al.* [13]. Other key principles that can be distilled from both guidance documents, playing in favor of an adaptive trial's external validity, are: prespecification (of the adaptations, decision rules, procedures, flow of information and so forth); parsimony (of the adaptations, number of people involved and so forth); information protection (e.g., regarding interim analysis results); and preferably – but not necessarily – maintaining the blind.

Q In 2006, Burman and Sonesson stated that 'the statistical community should strive to reach a consensus on the requirements that should be posed on the use of flexible designs and the related inference' [14]. A total of 5 years and two guidelines later, have we come closer to such a consensus? What gaps still remain?

MV: I think we have come closer to a consensus on the 'protective principles' for an adaptive trial: prespecification, parsimony, information protection and (preferably) maintaining the blind; as discussed earlier. Still, the crucial question of how flexible a trial should be may not be answerable in general terms. For example, are there really no circumstances where a confirmatory trial could be 'rescued' by an unplanned adaptation? I am not sure if a definite and general answer is possible. A guiding thought may be that any gain in flexibility should not come at substantial costs in terms of the feasibility, validity, integrity or efficiency of a trial [15].

AK: I think that the message in the executive summary of the European Reflection Paper is still of relevance: the importance of speeding up (or optimizing) drug development is well understood. The challenge, however, is to arrive there without lowering established standards for drug licensing or, even better, to improve the basis for (regulatory) decision making. The discussion between statisticians working in industry, academia and drug regulation has substantially improved our understanding of the importance of preplanning, trial integrity issues and practical issues of trial conduct. And it has lead to relevant developments in methodology and in experimental design allowing for increased flexibility under controlled conditions.

FM: The two regulatory guidance documents facilitate in an excellent way the journey of the statistical community to a common understanding. Different views on certain topics certainly still exist. What is needed now is increased practical experience from these designs which will help to converge the different opinions in the future.

Disclaimer

The views expressed are those of the authors and should not be understood or quoted as being made on behalf of the European Medicines Agency or its scientific committees, or be construed to represent the views or policies of the US FDA.

Financial & competing interests disclosure

B Gaydos is an employee of Eli Lilly and Company (salary and stock); F Miller is an employee of AstraZeneca (salary and stock); M Vandemeulebroecke is an employee of Novartis (salary and stock). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed. No writing assistance was utilized in the production of this manuscript.

References

Papers of special note have been highlighted as:

- ▪ of interest

1 European Medicine Agency. Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design. *European Medicines Agency* (2007).

- ▪ **Original European Medicine Agency reflection paper.**

2 US FDA. Adaptive design clinical trials for drugs and biologics. Draft guidance for industry. *US FDA* (2010).

- ▪ **Original US FDA draft guidance.**

3 *J. Biopharm. Stat.* 20, 1083–1219 (2010).

- ▪ **Comprehensive discussion of the European Medicine Agency reflection paper and the FDA draft guidance.**

4 ICH harmonised tripartite guideline E9. Statistical principles for clinical trials. *International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use* (1998).

5 ICH harmonised tripartite guideline E4. Dose-response information to support drug registration. *International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use* (1994).

6 Wang SJ, Hung HMJ, O'Neill R. Regulatory perspectives on multiplicity in adaptive design clinical trials throughout a drug development program. *J. Biopharm. Stat.* 21, 846–859 (2011).

7 Bornkamp B, Bretz F, Dmitrienko A *et al.* Innovative approaches for designing and analyzing adaptive dose-ranging trials. *J. Biopharm. Stat.* 17, 965–1032 (2007).

8 Ando Y, Hirakawa A. Discussion of: 'adaptive and model-based dose-ranging trials: quantitative evaluation and recommendations.' *Stat. Biopharm. Res.* 2, 462–465 (2010).

9 Pinheiro J, Sax F, Antonijevic Z *et al.* Rejoinder to: 'adaptive and model-based dose-ranging trials: quantitative evaluation and recommendations.' *Stat. Biopharm. Res.* 2, 455–468 (2010).

10 Hung HMJ, Wang SJ, O'Neill R. Flexible design clinical trial methodology in regulatory applications. *Stat. Med.* 30, 1519–1527 (2011).

11 Wang SJ. Commentary on: 'experiences in model/simulation for early phase or late phase study planning aimed to learn key design elements.' *Stat. Biopharm. Res.* 1(4), 462–467 (2009).

12 O'Neill R. FDA's critical path initiative: a perspective on contributions of biostatistics. *Biom. J.* 48, 559–564 (2006).

13 Brannath W, Burger U, Glimm E, Stallard N, Vandemeulebroecke M, Wassmer G. Comments on the draft guidance on 'adaptive design clinical trials for drugs and biologics' of the US Food and Drug Administration. *J. Biopharm. Stat.* 20, 1125–1131 (2010).

14 Burman CF, Sonesson C. Are flexible designs sound? (with discussion). *Biometrics* 62, 664–683 (2006).

15 Vandemeulebroecke M. Group sequential and adaptive designs – a review of basic concepts and points of discussion. *Biom. J.* 50, 541–557 (2008).