

Advancing clinical trial design in pulmonary hypertension

Andy P. Grieve¹, Shien-Chung Chow², John Curram³, Stephen Dawe⁴, Lutz O. Harnisch⁵, Noreen R. Henig⁶, Hsien Ming J. Hung⁷, D. Dunbar Ivy⁸, Steven M. Kawut⁹, Mohammad H. Rahbar¹⁰, Shen Xiao¹¹, and Martin R. Wilkins¹²

¹Aptiv Solutions, Innovation Centre, Stevenage Bioscience Catalyst, Gunnels Wood Road, Stevenage, Hertfordshire, SG1 2FX, UK, ²Duke University School of Medicine, Durham, NC, USA, ³Bayer PLC, Bayer House, Strawberry Hill, Newbury, RG14 1JA, UK, ⁴Novartis Pharmaceuticals UK Limited, Wimblehurst Road, Horsham, West Sussex RH12 5AB, UK, ⁵Pfizer, Pharmacometrics, Global Clinical Pharmacology, Sandwich, UK, ⁶Medical Affairs, Gilead Sciences, Inc. 333 Lakeside Dr, Foster City, California 94404, USA, ⁷DBI/OB/OTS/CDER/FDA, Building 21, Room 4616, 10903 New Hampshire Ave, Silver Spring, MD 20903, USA, ⁸Department of Pediatrics, University of Colorado, Pediatric Cardiology, Children's Hospital Colorado, USA, ⁹Department of Medicine and the Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia USA, ¹⁰The University of Texas School of Public Health at Houston, Center for Clinical and Translational Sciences, The University of Texas Health Science Center at Houston, 6410 Fannin Street, Houston, TX 77030, USA, ¹¹DCRP/CDER/FDA, Building 22, Room 4240, 10903 New Hampshire Ave, Silver Spring, MD 20903, USA, ¹²Experimental Medicine, Imperial College London, W12 0NN, UK

ABSTRACT

In pulmonary hypertension, as in many other diseases, there is a need for a smarter approach to evaluating new treatments. The traditional randomized controlled trial has served medical science well, but constrains the development of treatments for rare diseases. A workshop was established to consider alternative clinical trial designs in pulmonary hypertension and here discusses their merits, limitations and challenges to implementation of novel approaches.

Key Words: pulmonary arterial hypertension; clinical trial design; adaptive designs; modeling and simulation; Bayesian modeling; group sequential designs; population enrichment

An important obstacle to drug development for pulmonary arterial hypertension (PAH), as is the case with orphan diseases in general, is the ability to recruit a large enough sample of patients to draw inferences about a new drug's efficacy and safety. For most accepted intermediate or ultimate end-points, the sample size required to detect effects of clinical relevance is likely to be large relative to the available population. While there is no internationally accepted definition of an orphan, or rare, disease, the definitions in the EU, US and Japan span estimated prevalence of between four and seven cases per 10,000 individuals. These figures indicate that sample size will always be a critical issue.

A second issue is the mode of action of the new generation of therapies for PAH. If they are capable of effecting structural

change, then existing clinical trial designs may not be appropriate, or adequate, to detect such changes. Many study designs that have been used are appropriate for short-term alleviation of symptoms, but may be inappropriate for long-term outcomes, particularly when the effect of a drug diminishes over time. And thirdly, proving efficacy of new therapies added to background therapy is a challenge as it is hard to standardize background therapies or even stratify by background therapy.

Maintenance of a randomized treatment allocation is important to make inferences on comparisons of treatment differences. If the randomized dose allocation cannot be guaranteed during the conduct of the randomized part of a trial, inferences from the trial

Address correspondence to:

Dr. Martin R. Wilkins
Experimental Medicine
Imperial College London
Hammersmith Campus
Du Cane Road
London W12 0NN, UK
Email: m.wilkins@imperial.ac.uk

Access this article online

Quick Response Code:



Website: www.pulmonarycirculation.org

DOI: 10.4103/2045-8932.109933

How to cite this article: Grieve AP, Chow S, Curram J, Dawe S, Harnisch LO, Henig NR, et al. Advancing clinical trial design in pulmonary hypertension. *Pulm Circ* 2013;3:217-25.

can become questionable, that is, treatment changes during the randomized phase are likely confounding interpretation. In short-term trials (< 16 weeks), a randomization to placebo or another active treatment can be maintained, but the longer a trial runs, the greater the difficulty in maintaining the randomized treatments because of disease progression or because other confounding factors are changed (changes to standard of care, or investigators' judgments on treatment benefit). To assess disease progression, or the impact a potential treatment has on outcome over a longer period of time, for example, above one year, alternative treatments or therapies (hospitalization, lung transplantation, IV/SC epoprostenol) are often introduced which might affect the observation of the end-point. If mortality is the ultimate end point, it is likely that alternative treatments are introduced to varying degrees earlier or later in the trial, which can blur the causal relationships on treatment outcome and may render the interpretation difficult. The argument can be made that the introduction of additional or alternative treatments reflects everyday practice in treating real patients and increases the external validity of trials,^[1] and the issues here reflect the distinction between explanatory and pragmatic trials.^[2] The question of study design is pertinent to each of these issues and modern developments in clinical trial design should be considered alongside traditional designs.

RECENT REVIEWS OF CLINICAL TRIAL DESIGNS FOR PULMONARY ARTERIAL HYPERTENSION

There have been a number of reviews of appropriate clinical trial designs for PAH in the last few years.^[3-6] These reviews have largely covered similar designs and issues including (a) the use of a placebo control with background therapy, (b) noninferiority and equivalence trials, (c) randomized withdrawal designs, and (d) designs for combination therapies.

Use of a placebo control in untreated patients

Although the majority of approved therapies for PAH were tested in placebo-controlled studies in the absence of background therapy,^[6] there may be ethical issues around this approach going forward.^[7]

Noninferiority trials

If placebo-controlled studies are ethically controversial, an alternative is to consider an active control and to utilize a noninferiority hypothesis in which the experimental therapy is compared to an active control instead of a superiority hypothesis comparing the experimental medicine to placebo (superiority comparing experimental

therapy to a known therapy is a theoretic option but of questionable feasibility in the PAH population). What concerns researchers is that the sample size required may be larger than a placebo-controlled study and that the experimental conditions need to be "identical" to those that were in place when the comparator was licensed. This could be restrictive if there are developments in end-point technology since the efficacy of the comparator may not have been assessed on newer end-points. There are a number of other concerns surrounding noninferiority designs. First, patient heterogeneity increases the numbers required further. Second, it is easier to prove noninferiority against a drug with the same or very similar mechanism of action or ones with well-defined treatment effects and therefore it is a less desirable study design for a drug with a new mechanism of action (e.g., a drug which acts to improve right ventricular function compared with a pulmonary vasodilator). Third, for statisticians and regulators, the choice of the noninferiority margin is also of concern. Finally, there are concerns about assay sensitivity. Assay sensitivity is the ability of a trial to distinguish an effective treatment from a less effective or ineffective intervention. Without assay sensitivity, a trial is not internally valid and should not be used to compare the efficacy of two interventions because in such trials the lack of assay sensitivity may result in a conclusion that an ineffective intervention is noninferior leading to a false conclusion of efficacy.

Randomized withdrawal designs

Experts in PAH have also raised ethical issues concerning randomized withdrawal designs from two perspectives. First, to withdraw an effective therapy from a patient and to switch him/her to placebo would expose them to the chance that they may deteriorate and violate the necessary provision of the "standard of care" for the trial to be ethical. Second, even if a patient who deteriorates is withdrawn and given active therapy, there is no guarantee that they will return to their prerandomization state.

Designs for combination therapies

Factorial designs are in general underutilized in clinical research and would seem ideal for studying combination therapies. Given two drugs, A and B, factorial designs would comprise four groups:

- Placebo A + Placebo B
- Active A + Placebo B
- Placebo A + Active B
- Active A + Active B

Such a design is statistically optimal in that it would allow the estimation of the effect of each drug separately and their interaction, either synergistic or antagonistic effects. However, the use of the double placebo may create ethical issues as outlined above. There is one published factorial

trial in PAH which successfully used this approach in patients treated with background therapy.^[8]

If the double placebo is ethically unacceptable in a treatment-naïve population, two approaches may be possible. First, patients could be randomized to either A + B or B, where B is the current best standard of care, a randomized add-on design. Alternatively, a randomized withdrawal trial could be used in which patients who respond to A + B are randomized to either A + B or B, although this too may raise ethical concerns.

Randomized switch designs

An alternative to randomized withdrawal designs is the randomized switch design, in which one drug is withdrawn and another drug is substituted. The literature contains one example of this design in which patients were transitioned from IV epoprostenol to SC treprostinil or placebo over a period of up to 14 days.^[9] This study took measures to minimize ethical concerns about deterioration and nonetheless was able to establish a statistical difference between the randomized groups based on only 22 patients. Seven of eight patients (88%) withdrawn to placebo had clinical deterioration, while only one of 14 patients withdrawn to SC treprostinil deteriorated. It is perhaps relevant that this approach has not been repeated in more recent studies, which may have to do with a change in the ethical climate.

Adaptive designs

The term “adaptive design” refers to a clinical trial in which data collected during the course of the trial are used to change aspects of the trial design in such a way as to maintain the validity and integrity of the trial. The internal validity of a trial refers to the provision of correct statistical inference (e.g., appropriately adjusted *P* values and control of type I error) consistency between different stages of the trial and minimization of operational biases. The integrity of the trial refers to its ability to provide convincing evidence of an experimental drug’s efficacy, or lack of it, to a wider medical community and the maintenance of confidentiality of data during the running of the trial.

There are a number of objections to the use of adaptive designs in general, not confined specifically to PAH.^[5] These objections are statistical, methodological and ethical. One problem in the context of PAH is that the studies are generally so small that by the time sufficient data is collected to guide design, the trial is almost completely recruited. That said, the difficulty in recruiting patients to some developments would make them ideal for an adaptive design. In the context of pharmaceutical drug development, any adaptive design will need to meet the standards laid down in the guidelines issued by the regulators.^[10,11] The requirements of these guidelines address some of the issues which have been raised.^[5]

It is important to note that while some adaptive approaches allow almost total flexibility in changing many aspects of a trial during its course, in practice the defining philosophy of adaptive designs in the pharmaceutical industry is that they are adaptive by design. That is to say, those aspects of the trial that are open to adaptation should be prespecified in the study protocol and the consequences of these changes should be investigated through a comprehensive set of simulations prior to the commencement of a study.

Review of trials for pulmonary arterial hypertension in clinicaltrials.gov

To find out what type of trial designs have been used over the last decade, a search was made in the database www.clinicaltrials.gov using the search terms “PAH,” “Pulmonary Hypertension,” and “Pulmonary Arterial Hypertension.” Between 1993 and 2011, 201 studies were identified in the database. Figure 1 shows the start date of these studies, three-quarters of which began in 2001.

Figure 2 gives details of the characteristics of these studies. The studies were predominantly conducted in adults, in Phase III or Phase IV and were funded by industry. Approximately one-third of the trials were nonrandomized, a third of them were uncontrolled, approximately 40% were conducted as single-arm trials and over 50% were open label. Given the emphasis these days on the status of randomized, double-blind placebo-controlled trials, there is a surprising number of studies which are open-label, noncomparative and single arm.

RECOMMENDATIONS FOR EXPLORATORY PHASE OF DRUG DEVELOPMENT

Modeling and simulation for dose and dose regimen selection

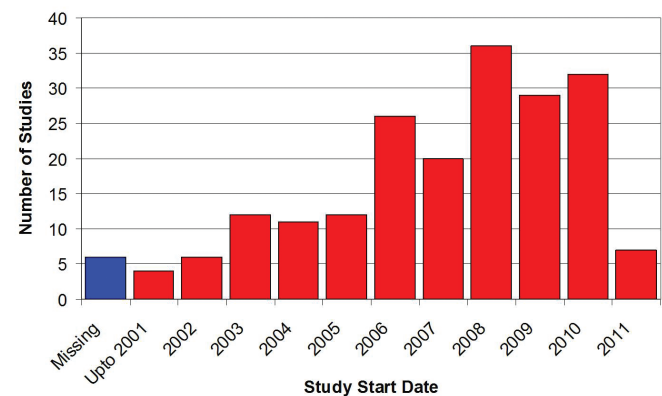


Figure 1: Start of pulmonary arterial hypertension trials (ClinicalTrials.gov: Extracted January 2011).

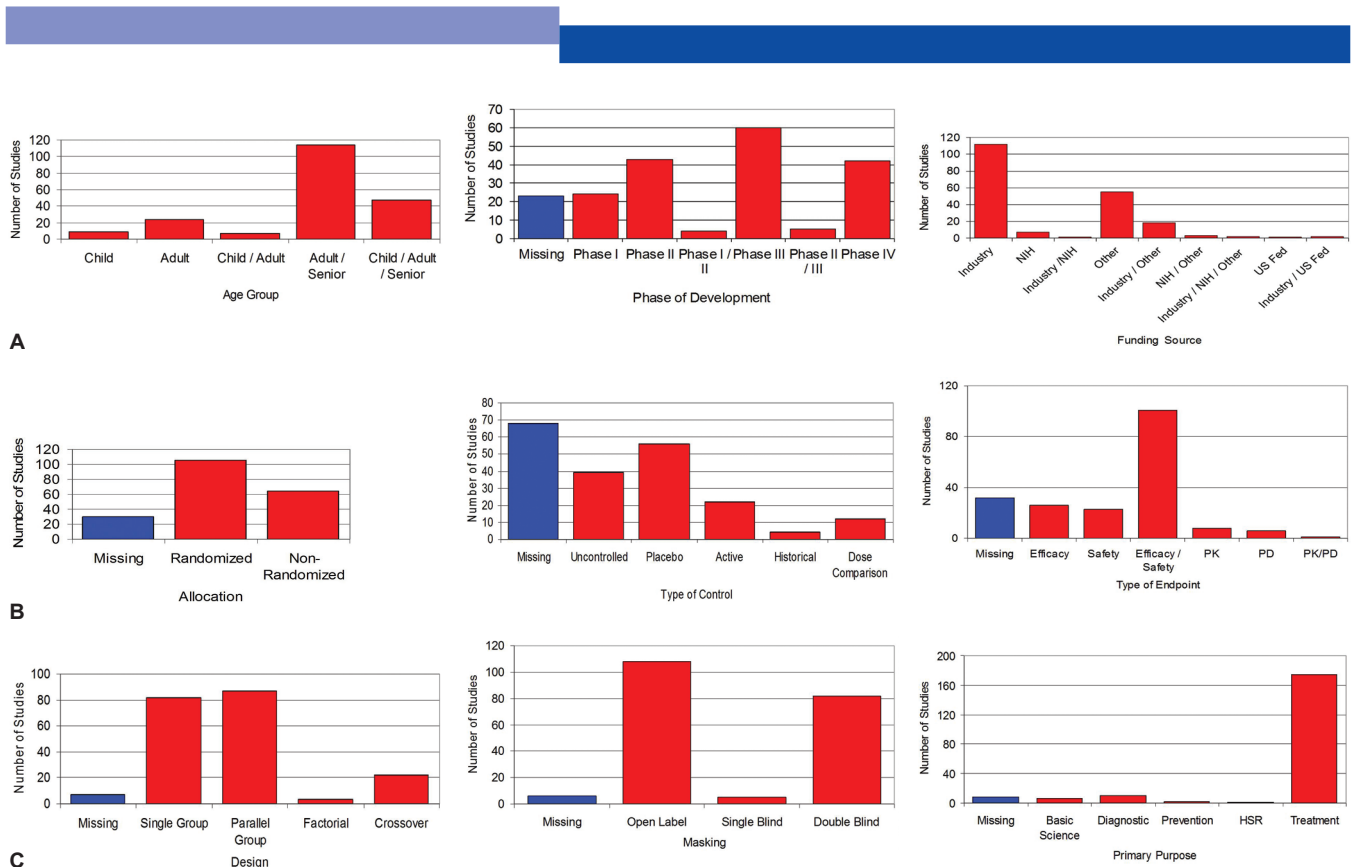


Figure 2: Study characteristics (ClinicalTrials.gov: Extracted January 2011). (A) Age, phase and funding. (B) Allocation, control and primary end-point. (C) Design, masking and primary purpose.

The incorporation of modeling and simulation techniques in drug development can establish linkages between underlying assumptions of clinical benefit, clinical efficacy and safety outcome measures and pharmacokinetic/pharmacodynamic data, for example, biomarker activity and blood pressure. In model-based drug development, physiological/statistical models of drug efficacy and safety can be developed based on all the available preclinical and clinical data, combining it with knowledge of disease progression and patient characteristics. Pharmacokinetic/Pharmacodynamic (PK/PD) models allow simulation of clinical trial scenarios to identify efficient and effective dose/dose regimens for future studies. Moreover, descriptive PK/PD modeling approaches can be applied to evaluate evidence of efficacy across the studied doses. In general, modeling offers a quantitative approach to improve drug development knowledge management and supports development of decision-making processes.^[12,13]

Bayesian modeling combined with use of external data to improve efficacy and safety signal detection in early development

As long ago as 1976, methods were introduced for combining historical control information with concurrent controls,^[14] and more recently, the advantages of this approach have been reiterated.^[15,16] Clearly when patient numbers are scarce, the ability to utilize historical control

information to reduce the number of patients randomized to the control is intuitively appealing. From a statistical perspective, care has to be taken to appropriately model such historical data taking account of study to study variation in control rates or control means. Hierarchical models, in particular Bayesian hierarchical models, may provide the appropriate framework for such modeling and the use of prior distributions about study to study variation is an essential element of this modeling. Examples of use of this approach are beginning to appear in the medical literature and its use for PAH should be considered.

Adaptive trial designs in early development

The development of Bayesian approaches to the design and analysis of studies in the exploratory phase of drug development has accelerated over the last decade. These approaches have utility in both traditional Phase I studies as well as Phase IIb dose selection studies.

In Phase I, the continual reassessment method (CRM) was originally developed to determine the Maximum Tolerated Dose (MTD) in oncology studies,^[17] but has subsequently been proposed for determining the Minimum Effective Dose (MED) in non-oncological applications.^[18] A criticism of the original CRM was that it too rapidly reached high doses, thereby increasing the risk that patients would suffer from dose-limiting toxicities. Various solutions to this problem

have been considered including escalation with overdose control^[19] and the probability interval approach.^[20]

In Phase IIb studies, a complex Bayesian approach to dose selection has been proposed.^[21] The concrete example for which the methodology was developed^[22] involved 15 doses of an experimental drug, but has also been used for cases with fewer doses.^[23]

The argument in favor of such adaptive dose-ranging trials is that they are more efficient than traditional approaches. There are a number of reasons for this. First, such studies are designed to explicitly address the goals of the individual trial through choice of appropriate decision criteria. Second, the approach to analysis in such designs recognizes that in the learning phase of drug development,^[24] it is more appropriate to estimate the dose–response function rather than to compare individual doses to placebo. Finally, the design reduces the chance of having to re-run a Phase IIb design by a poor choice of a few doses.

RECOMMENDATIONS FOR CONFIRMATORY PHASE OF DRUG DEVELOPMENT

Sample size re-estimation within a confirmatory trial (Phase III)

The conundrum of sample sizing is that it is inherently based on an a priori “best guess.” The ideal sample size estimate would require information which, if precisely known, would cast doubt on the need for the study. An erroneous sample size estimate may lead to a study which is too small and therefore underpowered, or too large, entailing increased expense and potentially unnecessary risks and burdens to more subjects than necessary. Therefore, estimating the optimal sample size is essential. It has been argued that the main benefit of an adaptive approach in drug development comes from its ability to stage the utilization of resources and (in essence) the risks and burdens to patients.^[25] Such trials begin with a relatively small sample size, with additional resources and subjects being utilized only if necessary or if the results are “promising.”

Recent developments in group sequential designs

Although group sequential designs (GSDs) have been around for 35 years,^[26] their use has not been as widespread as might originally have been envisaged. The principal idea of GSDs is to allow sponsors to examine the accruing data in a trial on a number of occasions, interims, and to stop the trial if there is statistically significant evidence that the experimental drug is effective, or if there is no evidence of its efficacy – so-called futility. The

advantages are clear. In the former case, the decision as to the efficacy of an experimental drug may be made with reduced numbers of patients and expedite the dissemination of the information and potentially drug approval, while in the latter, if there is no evidence of the drug’s efficacy on both resource and ethical grounds, it is beneficial to stop the trial. Stopping early for efficacy may raise issues of the amount of evidence of an experimental drug’s safety. While this can be overcome by post-licensing studies and/or conditional approval processes, it is not easy to withdraw a drug once approved. The downside is the potential for increasing the false-positive rate by multiple looks at the data, but they were developed precisely to control the overall type-I error and there is an inevitable concentration on the primary end-point with a consequent loss of information on important secondary end-points.^[26,27] One aspect of GSDs that has recently been investigated is the issue of long-term outcomes and loss of efficiency. If the primary end-point of a study is long term, say one year, then at interim analysis many patients may have been entered into the study and not yet have reached the one-year time point. Therefore, in assessing the significance of treatment effects and making decisions about whether the trial should stop or not, many patients who have been randomized will not contribute to the decision-making process, jeopardizing the ethical justification for their enrollment (the altruistic contribution to answering a scientific question).

A number of approaches have been proposed to overcome this issue. If early (continuous) measurements are available for each patient, then this early information can be used to increase the efficiency of decision making in GSDs.^[28] Similar proposals have been proposed for binary data.^[29,30] In a complex Bayesian adaptive Phase IIb design, in a “longitudinal model,” predictions of final outcomes are combined with real final outcomes in updating a Bayesian prior of the dose–response curve^[21] which is closely related to the previous approach.^[28] More recently, a class of GSDs has been developed which takes account of what are called “pipeline patients” and, by doing so, reduces the risk that a decision to stop based on only observed data may be “overtaken” when the complete data is available on all patients.^[31]

Adaptive trials in a confirmatory setting

Methodology has been developed which allows for extreme flexibility in running adaptive trials, even in a confirmatory setting.^[32] The methodology can accommodate changes to end-point and sample size, dropping of arms, changing of test statistics, modifications to the population, treatment duration, patient population, number of treatments, number of interim analyses, hypotheses and the combination of multiple study objectives in a single trial.

Seamless Phase II/III designs are perhaps the most notable examples of adaptive designs in a confirmatory setting. They are considered to be “one of the main types of adaptive approaches used currently in confirmatory settings” which “add efficiency in the use of patient data, draw stronger conclusions with the same number of exposed patients and reduce development time.”^[33] Seamless Phase II/III trials are carried out in two parts. The first part is effectively a Phase IIb exploratory dose-response study, in which experimental treatments (or doses) are compared to a control treatment. The aim is to determine the most promising treatment (dose) or treatments (doses). Treatments (doses) which show sufficient promise are retained for the second part – effectively a Phase III, confirmatory trial – together with the control treatment. At the conclusion of the trial, data from both parts are combined in the final analysis to assess the efficacy of the selected treatment(s). Flexible methodology is available to preserve the validity of such a trial by strong control of the family-wise type-I error rate.^[34]

Seamless Phase II/III clinical trials are more efficient than separate Phase II and Phase III programs in that fewer patients are required to achieve a given program-level power. The benefit arises from the inclusion of the first stage data on the selected treatment, suitably adjusted for multiplicity, in the final analysis at the end of the trial. To illustrate this latter point, Figure 3 shows a comparison between the total sample sizes in an inferentially seamless Phase II/III study in which the Phase II part began with four doses and at the end of Phase II, a single dose was selected for the Phase III part and what is often termed an operationally seamless Phase II/III study in which the Phase III part is independently sample sized. These results arose in designing a Phase II/III seamless program for an orphan disease with binary end-point simulations which were run on 20 different dose response scenarios. Figure 3 shows a consistent increase if the study is run in an operationally seamless fashion of approximately 1/3. For an orphan disease, these savings in sample size by running the trial inferentially seamlessly are extremely important.

Of course we need to be aware of caveats to the use of seamless Phase II/III designs. First, since all adaptations are prespecified, occasional “learnings” in Phase II could not lead to adjustments for the Phase III design, which could lead to potential Phase III failure. This risk would always need to be weighed against the perceived benefits. Second, the primary end-point for confirmation is prespecified and will be measured on all patients. Third, there is positive data from Proof-of-concept (POC) studies with the remaining uncertainty primarily concerning the dose. Fourth, the marketing

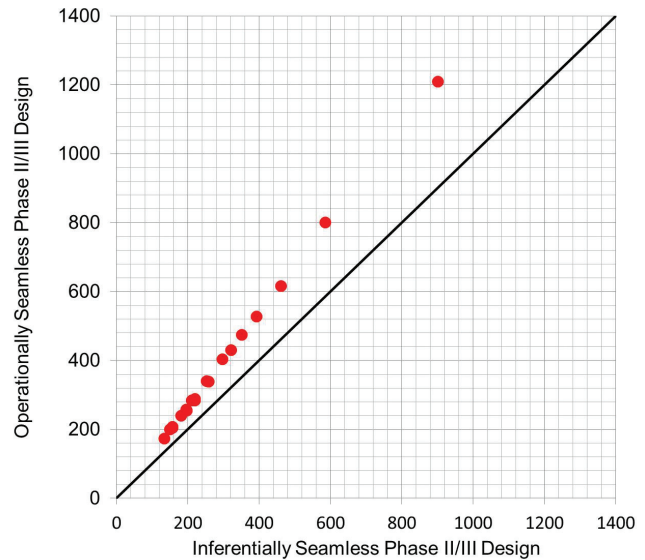


Figure 3: Comparison of total sample sizes of inferentially and operationally seamless Phase II/III design.

formulation of the test drug is available. Finally, the patient population is defined and will stay the same in both phases of the study.

POPULATION ENRICHMENT AND SUBGROUP ANALYSIS

A biomarker has been described as “... a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention.”^[35] Biomarkers can play many different roles in drug development, including a predictor of response or resistance to specific therapies, being a correlative end-point and used in longitudinal models referred (*see* “Recent developments in group sequential designs”), being a surrogate end point in modeling and simulation (*see* “Modeling and simulation for dose and dose regimen selection”), or being used as a means for patient-enrichment designs.

Factors used to limit the study population to patients believed more likely to benefit from the experimental therapy are termed enrichment factors.^[36,37] Such factors may be predictive biomarkers, or they may be biomarkers, clinical-pathologic, demographic characteristics associated with a predictive biomarker or with the target of a therapeutic agent. The smaller the proportion of truly benefiting patients in the population, the more advantageous it is to consider studying an enriched population.

Biomarkers that could be useful as enrichment factors during the drug development process might still need further refinement before they are ready for clinical use

as predictive factors. This is because many enrichment biomarkers used in drug development either do not have sufficiently high positive or negative predictive value to justify clinical use or the assay used to measure the biomarker during the drug development process might not be sufficiently robust and reproducible for routine clinical use. The main purpose of using an enrichment biomarker in drug development is to improve the chances that the drug will show benefit in the tested subgroup of patients to more quickly establish that the drug is worth pursuing further. If information is available to suggest subgroups of patients who are more likely to benefit from a therapy, it may be reasonable to conduct a confirmatory trial only in those patients.

These kinds of observations are growing in medicine, where increasing use of molecular signatures reveals that the traditional tools used for diagnosis are lumping diverse phenotypes together. A recent report calls for precision medicine by which is meant the use of genomic, epigenomic, exposure and other data to define individual patterns of disease and phenotypes with more granularity, potentially leading to better individual treatment.^[38] Precision medicine couples established clinical-pathologic indexes with state-of-the-art molecular profiling to create diagnostic, prognostic and therapeutic strategies tailored for specific groups of patients.

The aspect of “one size fits all” surrounding the conventional design of clinical trials has been challenged, particularly when the diseases are heterogeneous due to observable clinical characteristics and/or unobservable underlying genomic and epigenomic characteristics and/or the experimental therapy is tailored to specific mechanism of action. An extension from the traditional single population design objective to one in which several possible patient subpopulations are studied will allow more informative evaluation in the patients having different degrees of responsiveness to the therapy. Building into traditional clinical trials a prospectively planned selection of subpopulations with higher response to the therapy is appealing from the patient’s perspective as it addresses personalized medicine in adequate and well-controlled clinical trials. These new adaptive designs, called adaptive patient-enrichment or population-enrichment designs, allow modification to study hypothesis, the reallocation of patients and re-estimation of the sample size midstream to achieve the pre-planned objective.

It has been shown recently that such adaptive enrichment designs can be constructed to study a clinical hypothesis of treatment effect in the full population as well as several hypotheses of treatment effect in prespecified subsets more efficiently than the conventional nonadaptive approach.^[39-41] The statistical methodology is very similar to

the statistical methodology of seamless Phase II/III designs referred to above.

While in a seamless Phase II/III design the adaptation relates to the selection of treatment arms, in the enrichment design the primary selection concerns the population. Such a study progresses seamlessly either in the subpopulation(s) of patients or in the whole population on the basis of data obtained in the first stage. At the end of the trial, the data from both stages are combined in the final analysis to assess the efficacy of the selected subpopulation(s), preserving its validity by strong control of the family-wise type-I error rate. As in the seamless Phase II/III design, enrichment designs may be more efficient than separate Phase II and Phase III programs in that fewer patients are required to achieve a given program-level power. Again, the benefit arises from the inclusion of the final stage data on the selected subpopulations, suitably adjusted for multiplicity, in the final analysis at the end of the trial.

An example of a complex population-enrichment design is the ongoing I-SPY 2 TRIAL (Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging And moLecular Analysis) trial involving a randomized Phase II stage in which a number of experimental agents are tested in combination with standard neoadjuvant chemotherapy for patients with high-risk primary breast cancer.^[42] The primary end-point is pathologic complete response (pCR) at the time of surgery, with the objective being to identify biomarker signatures that predict pCR for drugs or combinations of drugs. The study is to be used to evaluate many drugs and drug/biomarker combinations, with successful combinations being “graduated” to a Phase III study and failures being dropped for futility.

One particular aspect of the I-SPY 2 TRIAL trial is the co-operative nature of the study in that multiple sponsors provided the experimental agents that are being used. The advantage to individual sponsors is to spread the costs by the use of a single control group. But the absence of a definitive noninvasive biomarker would hamper this approach in PAH.

Whether these types of designs will be used in PAH remains to be seen. Given the orphan nature of the disease and the difficulty of recruitment, it may be optimistic to expect that sufficient patients will be available to conduct such subgroup searches. Certainly we may need to restrict the number of subgroups considered.

DISCUSSION

The traditional clinical trial has been with us for over 60 years and has served us well. Nevertheless, there

are reasons for believing that there is a need to consider other ways of developing new drugs.

The last 10 to 15 years have seen considerable interest in adaptive clinical trial designs in the hope that they can contribute to improving the success rate particularly in Phase III clinical trial programs. But if adaptive clinical trials are to be more widely used, there are a number of issues that need to be kept in mind.

First, an adaptive design cannot make a noneffective drug effective. The purpose of adaptive designs is to increase the likelihood of detecting an effective drug efficiently.

Second, the use of adaptive designs is not an excuse for poor planning. If anything, the planning of adaptive designs takes more time and effort than nonadaptive designs. But this additional time is worth it since, even if in the planning phase it is decided not to run a trial adaptively, the extra planning time will ultimately make the clinical trial run better than it would have without the extra thought.

Third, adaptive designs are not a charter for protocol amendments. Adaptive designs are adaptive by design. That is, they are planned to be adaptive from the outset and thought is given to those aspects of the trial design which have potentials for adaptation.

Fourth, while clinical trial methodologists have, with considerable ingenuity, broadened the scope for adaptive designs, the judicious use of a few adaptations is preferable to trying to include as many adaptations as is feasible. This is not only on practical grounds, but also a regulatory imperative, particularly in the confirmatory phase of drug development. The argument made by regulators is that in a confirmatory trial if there are a large number of adaptations, in what sense can it be regarded as confirmatory.

Finally, the implementation of adaptive clinical trials can be more complicated than traditional trials and getting the implementation right is crucial to the success of an adaptive design.

Alternative trial designs have also been discussed, in particular the enrichment and subgroup analyses. If the particular subpopulation at risk or likely to respond to treatment could be identified, or if a promising set of biomarkers were to show sufficient predictability, it might lead to a more experimental, assumption-rich clinical trial setting in which the focus would be on the exploratory rather than the confirmatory nature of the development program.

Using model-based approaches which provide a methodology to integrate the various sources of

preclinical and clinical information in a quantitative way, a framework can be established to capture the assumption-rich situation prevalent in rare diseases or special populations. The development of such a knowledge base allows for the building of predictive models for future trial outcomes or priors for Bayesian models, but might also allow for the substitution of evidence generation if the inferences from a model-based approach are perceived strong enough.

Recently some regulatory agencies have started discussions to develop guidance which could allow a consistent and transparent extrapolation of safety and efficacy measures in circumstances when sufficient evidence is difficult to gather, as for example in pulmonary hypertension and inferences from a modeling approach could potentially substitute for real data under those special circumstances.^[43]

CONCLUSIONS

There are considerable challenges to running randomized controlled trials (RCTs) in PAH and we have sought to set out and discuss, the designs that are available. While further methodological work needs to be carried out to define those designs which are particularly suited for PAH, a more adventurous approach needs to be taken by industry, regulators and academic researchers alike to improve the conduct and interpretability of trials.

As an orphan indication PAH has particular difficulties, not least of which is that recruitment into trials is usually slow and therefore trials span a long period of time if markers for disease progression are being studied. These are exactly the circumstances in which adaptive designs can be considered.

The judicious choice of innovative designs, despite the obstacles, should be considered in order to raise the likelihood of new treatments reaching patients.

Steering Committee for Working Group 3

Shein-Chung Chow (sheinchung.chow@duke.edu),
John Curram (john.curram@bayer.com),
Stephen Dawe (stephen.dawe@novartis.com),
Andy Grieve (Andy.Grieve@aptivsolutions.com) – Principal author,
Lutz O Harnisch (lutz.o.harnisch@pfizer.com),
Noreen Henig (Noreen.henig@gilead.com),
H. M. James Hung (hsienming.hung@fda.hhs.gov),
D. Dunbar Ivy (Dunbar.Ivy@childrenscolorado.org),
Steven M. Kawut (kawut@mail.med.upenn.edu),
Mohammad H. Rahbar (Mohammad.H.Rahbar@uth.tmc.edu),
Martin Wilkins (m.wilkins@imperial.ac.uk) – Chair,
Shen Xiao (shen.xiao@fda.hhs.gov),

REFERENCES

1. Rothwell PM. External validity of randomised controlled trials: "To whom do the results of this trial apply?". *Lancet* 2005;365:82-93.
2. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis* 1967;20:637-48.
3. Hooper MM, Oudiz RJ, Peacock A, Tapson VF, Howarth SG, Frost AE, et al. Endpoints and clinical trial designs in pulmonary arterial hypertension. *J Am Coll Cardiol* 2004;43:48S-55.
4. Gombert-Maitland M. Traditional and alternative designs for pulmonary arterial hypertension trials. *Proc Am Thorac Soc* 2008;5:610-6.
5. McLaughlin VV, Badesch DB, Delcroix M, Fleming TR, Gaine SP, Galie N, et al. End points and clinical trial designs in pulmonary arterial hypertension. *J Am Coll Cardiol* 2009;54:S97-107.
6. Peacock AJ, Naeije R, Galie N, Rubin L. End-points and clinical trial designs in pulmonary arterial hypertension: Have we made progress? *Eur Respir J* 2009;34:231-42.
7. Halpern S, Doyle R, Kawut SM. The Ethics of Randomized Clinical Trials in Pulmonary Arterial Hypertension. *Proc Am Thorac Soc* 2008;5:631-5.
8. Kawut SM, Bagiella E, Lederer DJ, Shimbo D, Horn EM, Roberts KE, et al. Randomized clinical trial of aspirin and simvastatin for pulmonary arterial hypertension: ASA-STAT. *Circulation* 2011;123:2985-93.
9. Rubenfire M, McLaughlin VV, Allen RP, Elliott G, Park MH, Wade M, et al. Transition from IV epoprostenol to subcutaneous treprostinil in pulmonary arterial hypertension: A controlled trial. *Chest* 2007;132:757-63.
10. European Medicines Agency, 2009. Reflection Paper on Methodological Issues in Confirmatory Clinical Trials Planned With an Adaptive Design [Internet]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003616.pdf. [Last accessed 2012 Oct 01].
11. Food and Drug Administration, 2010. Adaptive Design Clinical Trials for Drugs and Biologics (Draft Guidance) [Internet]. Available from: <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm201790.pdf>. [Last accessed on 2012 Oct 01].
12. Miller R, Ewy W, Corrigan BW, Ouellet D, Hermann D, Kowalski KG, et al. How modeling and simulation have enhanced decision making in new drug development. *J Pharmacokinet Pharmacodyn* 2005;32:185-97.
13. Lalonde RL, Kowalski KG, Huttmacher MM, Ewy W, Nichols PA, Milligan PA, et al. Model-based drug development. *Clin Pharmacol Ther* 2007;82:21-32.
14. Pocock SJ. The combination of randomized and historical controls in clinical trials. *J Chronic Dis* 1976;29:175-88.
15. Spiegelhalter DJ, Abrams KR, Myles JP. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Chichester: John Wiley and Sons; 2003.
16. Neuenschwander B, Capkun-Niggli G, Branson B, Spiegelhalter DJ. Summarizing historical information on controls in clinical trials. *Clin Trials* 2010;7:5-18.
17. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: A practical design for phase 1 clinical trials in cancer. *Biometrics* 1990;46:33-48.
18. Zohar S, Resche-Rigon M, Chevret S. Using the continual reassessment method to estimate the minimum effective dose in phase II dose-finding studies: A case study. *Clin Trials* 2011; published online August 8th 2011.
19. Babb J, Rogatko A, Zacks S. Cancer Phase I clinical trials: Efficient dose escalation with overdose control. *Stat Med* 2012;17:1103-20.
20. Neuenschwander B, Branson M, Gsponer G. Critical aspects of the Bayesian approach to Phase I cancer trials. *Stat Med* 2008;27:2420-39.
21. Berry DA, Müller P, Grieve AP, Smith MK, Parke T, Blazek R, et al. Adaptive Bayesian designs for dose-ranging trials. In: Carlin B, Carriquiry A, Gatsonis C, Gelman A, Kass RE, Verdine I, et al. editors. *Case Studies in Bayesian Statistics V*, New York: Springer Verlag; 2002. p. 99-181.
22. Grieve AP, Krams M. ASTIN: A Bayesian adaptive dose-response trial in acute stroke. *Clin Trials* 2005;2:340-51.
23. Berry SM, Spinelli W, Littman GS, Liang JZ, Fardipour P, Berry DA, et al. A Bayesian dose-finding trial with adaptive dose expansion to flexibly assess efficacy and safety of an investigational drug. *Clin Trials* 2010;7:121-35.
24. Sheiner LB. Learning versus confirming in clinical drug development. *Clin Pharmacol Ther* 1997;61:275-91.
25. Mehta CR, Pocock SJ. Adaptive increase in sample size when interim results are promising: A practical guide with examples. *Stat Med* 2012;30:3267-84.
26. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. *Biometrika* 1977;64:191-9.
27. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1977;35:549-56.
28. Galbraith S, Marschner I. Interim analysis of continuous long-term endpoints in clinical trials with longitudinal outcomes. *Stat Med* 2003;22:1787-805.
29. Marschner IC, Becker SL. Interim monitoring of clinical trials based on long-term binary endpoints. *Stat Med* 2001;20:177-92.
30. Sooriyachchi MR, Whitehead J, Whitehead A, Bolland K. The sequential analysis of binary responses: A score test for the case of three time points. *Stat Med* 2005;25:2196-214.
31. Hampson LV, Jennison C. Group sequential tests for delayed responses. *J R Stat Soc Series B* 2013;75:1-37.
32. Bauer P, Köhne K. Evaluation of experiments with adaptive interim analyses. *Biometrics* 1994;50:1029-41.
33. Gallo P, Anderson K, Chuang-Stein C, Dragalin V, Gaydos B, Krams M, et al. Viewpoints on the FDA draft adaptive designs guidance from the pharma working group. *J Biopharm Stat* 2010;20:1115-24.
34. Posch M, König F, Branson M, Brannath W, Dunger-Baldauf C, Bauer P. Testing and estimating in flexible group sequential designs with adaptive treatment selection. *Stat Med* 2005;24:3697-714.
35. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001;69:89-95.
36. Temple R. Special study designs: Early escape, enrichment, studies in non-responders. *Commun Stat Theory Methods* 1994;23:499-531.
37. Temple R. Enrichment designs: Efficiency in development of cancer treatments. *J Clin Oncol* 2005;23:4838-9.
38. National Academy of Sciences. *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease*. Washington DC: The National Academies Press; 2011.
39. Brannath W, Zube E, Branson M, Bretz F, Gallo P, Posch M, et al. Confirmatory adaptive designs with bayesian decision tools for a targeted therapy in oncology. *Stat Med* 2009;28:1445-63.
40. Wang SJ, Hung HMJ, O'Neill RT. Adaptive patient enrichment designs in therapeutic trials. *Biom J* 2009;51:358-74.
41. Wang SJ, O'Neill R, Hung J. Approaches to evaluation of treatment effect in randomized clinical trials with genomic subset. *Pharm Stat* 2007; 6:227-44.
42. Barker AD, Sigman CC, Kelloff GJ, Hylton NM, Berry DA, Esserman LJ. I-SPY 2: An adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clin Pharmacol Ther* 2009;86:97-100.
43. European Medicines Agency, 2012. Concept paper on extrapolation of efficacy and safety in medicine development. EMA/129698/2012 [Internet] (available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500129285.pdf) [Last accessed on 2012 Oct 01].

Source of Support: Nil, **Conflict of Interest:** None declared.