ADAPTIVE DESIGN YEARS LATER: What Have We Learned?

Designing and executing adaptive clinical trials effectively is hard. Teams attracted by the potential for more flexible trial designs to result in faster, more efficient clinical development programs have learnt these gains are mitigated by increased complexity and operational challenges. The successes and failures of these teams provide lessons for everyone. In this paper, we look at those lessons and how they have enabled the consistent, efficient design and execution of complex trials.

Researchers began discussing multi-stage adaptive design in the late 1980s. The goal then was to enable mid-trial modifications based on unblinded and external data without raising the risk of false positives. Some people criticized statistical aspects of adaptive design but the potential for it to make trials more flexible and efficient prompted others to pursue and refine the concept. This led to the development of regulatory guidance documents and the establishment of interim data reviews as an everyday aspect of drug and medical device clinical trials.

That top-line summary of the story of adaptive design masks innumerable advances and setbacks that litter the path it has taken from controversial idea to routine feature of clinical trials. Over its near-30-year history, adaptive design has passed through hype cycles as excitement about the potential of the approach has outpaced the ability to deliver on its promise. These shortcomings in delivery have resulted in more than just suboptimally designed and executed trials. Sponsors have run into serious problems because of flaws in the design and execution of adaptive trials.

The positive outcome of the setbacks is that leading teams now know exactly what it takes to succeed and have invested in the adaptive software, technology, experience, and experts they need to do so. These teams consistently design and execute adaptive trials that enable time and cost savings, optimized program development, and earlier, better-informed go/no-go decisions.

Lessons Learnt From Decades Of Implementing And Executing Adaptive Trials

The success of these teams is built on decades of hard-won lessons. Adaptive design is no longer a new, unproven concept. Each adaptive trial, whether it is a qualified or unmitigated success or failure, provides feedback on how to best design and execute the next study.

This has led to the identification of three critical operational challenges in adaptive clinical trials: data quality, operational bias and implementing adaptations, and sample size reassessments. The importance of these topics stems from their significance to regulatory agencies. Regulators need to be certain about the

outcomes of trials. Data quality issues, operational bias, and interpretability questions prevent such certainty and undermine the findings of trials.

The potential for poor-quality data to create uncertainty unique to adaptive trials is a result of the use of interim analyses. These analyses allow sponsors to stop trials that are clearly destined to succeed or fail early. In the case of studies with strong signs of efficacy, such stoppages cut the time it takes to bring an effective product to market and the broader patient population. Similarly, when a drug or device is clearly failing to improve outcomes, an early stoppage frees patients to leave the trial and start treatment with a more effective intervention.

These characteristics make interim analyses popular with sponsors and investigators. Yet, interim decisions are only as good as the data on which they are based. Data must be consistent and accurate when the interim analysis is performed. If a sponsor makes a mid-trial decision based on a dataset that is later shown to be compromised, the adaptive change and whole study may be invalidated. This means the study team must factor in all protocol deviations and other events that may affect the dataset before performing the interim analysis.

This is far from a hypothetical concern. One sponsor stopped a Phase III chronic graft-versus-host disease study early after a second interim review showed its efficacy measure just cleared the threshold of statistical significance. The riskiness of halting the trial based on marginal efficacy was revealed later when evidence of a major protocol deviation emerged. The deviation affected data from five patients. When the dataset was corrected, the efficacy measure fell short of the threshold of statistical significance. The trial failed.

This case study shows why data used in interim analyses must be complete and accurate. Teams running adaptive trials cannot wait until the end of the study to clean their data. Cleaning must be done often enough that teams can perform interim data reviews within a few days. This is best achieved by using risk-based technology to monitor and clean data in real time. The Phase III would also have benefited from using a software solution to run simulations and analyses to support robust interim decisions.

When paired to best practices such as the use of the same quality standard for interim and final database locks, technologies increase the chance of success. That also applies to the second and third critical challenges, operational bias and sample size reassessment.

Operational bias occurs when unblinded information or knowledge of adaptive decisions or rules leak to investigators. The leaking of unblinded information comprises data integrity. The release of information about adaptive decisions and rules can affect the behavior of investigators. For example, if investigators know the sponsor will stop poorly-performing arms of the study following interim reviews, they may delay enrollment to increase the likelihood of their patients receiving an effective treatment. This makes adaptive decisions a potential source of bias. All sources of bias are devastating. No statistical adjustments can undo the bias.

The sponsor of a Phase III clinical trial in acute infectious diarrhea learnt this the hard way. The data were well below the threshold for statistical significance when the first and second interim analyses were performed. At the second review, the target enrollment number was raised. This suggested efficacy was weak. Enrollment slowed following the decision, prompting the sponsor to halt the trial early. At the final analysis, the efficacy data cleared the newly-lowered threshold for success. Such quick reversals, from likely failure to ultimate success, following an adaptive change are red flags for operational bias. The results of the trial were therefore closely scrutinized.

Effective use of technology and best practices could have prevented the perception of bias. When designing and running trials with planned adaptive sample size re-estimation, it is advisable that study protocols only refer to the maximum sample size. That way, investigators cannot intuit when the sponsor increases the sample size. The sponsor can mitigate the risk investigators will learn of adaptive changes by limiting dissemination of decisions to a small number of its personnel. Finally, companies can use Interactive Randomization Technology to encrypt and restrict access to information about randomization and drug supply.

How Sophisticated Software, Technology, And Best Practices Are Improving Implementation

Leading adaptive design groups have responded to the lessons learnt in recent years by stepping up their investments in people, integrated technology, best practices, and software. The result is some groups are now experts in the design, implementation, and execution of complex adaptive trials across all phases of development.

Quick reversals, from likely failure to ultimate success, following an adaptive change are red flags for operational bias.

Software for designing, simulating, and analyzing adaptive clinical trials of all phases is the bedrock of these groups' success. Teams that possess such software apply it to adaptive group sequential designs and sample size re-estimation, adaptive multiple comparison procedures, adaptive population enrichment designs, and adaptive dose finding designs.

Taking dose finding as an example shows the power of such software. Identifying the maximum tolerated dose is a vital step in oncology drug development. Yet, until recently, the vast majority of trials used the simple rule-based 3+3 escalation scheme, despite evidence it underestimates the maximum dose.

The latest software makes it easier for sponsors to choose alternative, better options by simulating other approaches. Users can compare 3+3 designs to the continual reassessment method, Bayesian logistic regression with overdose control, and the modified toxicity probability interval approach. Having used one software module to select the best approach based on simulations, sponsors then turn to another for recommendations on the next and final target dose during the study.

Elite adaptive design teams are using these and other capabilities to design and run ever-more complex studies, such as cluster trials. These bring together the concepts of umbrella, basket, and platform trials. Cluster trials consist of sub-trials targeting specific cancer types, phenotypes, or genetic signatures. This ensures subpopulation homogeneity while also facilitating information sharing between sub-trials, making the approach a good fit for precision medicine and immuno-oncology combinations.

Today, work is underway to create scientific machinery that further saves resources, optimizes development, and enhances knowledge about the effects of drugs and devices.

One cluster trial is testing combinations of five drugs in six types of cancer. This creates complexity. Each pair of cancer types must be tested in two sub-trials. These sub-trials must enroll three cancer types and administer two drugs. Every drug must be tried in every cancer type in two sub-trials.

Only organizations with the staff and computing power to combine structured and individual level exploration and include operational models in the trial design can plan and execute such studies. Extensive use of Monte-Carlo simulations is needed to understand the statistical properties of the designs.

Organizations that can handle these tasks are rewarded with trials that enable the efficient testing of multiple combinations of drugs in different types of cancer. In the above-cited cluster trial, the CRO cut the number of sub-studies

from 60 to 30, enabling the sponsor to identify effective drug combinations faster and more efficiently.

Sponsors developing drugs in indications other than cancer also benefit from adaptive design. Take, for example, a rheumatoid arthritis trial that began with a multiple ascending dose stage. At that stage, an unblinded medical monitor and biostatistician made futility decisions for the dose cohorts based on ACR20 and the reduction in c-reactive protein after four weeks. The unblinded researchers performed 10 interim analyses and took no more than three days for each review. Such speed is only possible if data is cleaned in real time using technology

Once the maximum tolerated dose was reached, the trial moved seamlessly into its second stage. The design enabled the trial to enroll and randomize 253 patients in one year while realizing the efficiency benefits of combining Phase I and II. Running one trial rather than two makes regulatory review, site startup, and study close-out more efficient. In the rheumatoid arthritis trial, the savings amounted to about $1.2 million and nine months of development time.

Stakeholders other than sponsors benefit from such trials. The rheumatoid arthritis trial stopped the lowest-dose cohort early. That spared subjects from having to continue taking an ineffective regimen, a benefit that makes adaptive design popular with patients and the physicians who care for them.

Regulators and payers also advocate for better trial design. The US Food and Drug Administration (FDA) and European Medicines Agency (EMA) have both established positions on the use of adaptive design. This has raised familiarity with — and confidence in — the approach among payers, a group that benefits from well-designed adaptive clinical trials.

Traditional trials deliver data needed to win regulatory approval but struggle to also provide results that show the value of a drug to payers. This leaves sponsors needing to gather additional results to show the effect of the drug in the real world or identify the patients most likely to respond. The flexibility provided by adaptive design mitigates some of these challenges by improving the selection of doses and identification of biomarker-defined subpopulations of patients.

The Future Of Adaptive Design

The lessons, and responses to them, described in this paper have brought the industry to the point that teams can consistently and efficiently design and execute effective, complex adaptive clinical trials. That is not the end of the process, though. Today, work is underway to create scientific machinery that further saves resources, optimizes development, and enhances knowledge about the effects of drugs and devices.

Teams at the forefront of adaptive design are looking to technological breakthroughs to advance the field, for example by applying quantum computing to machine learning models to predict clinical trial outcomes. Such initiatives will redefine how sponsors use data and adaptive design to accelerate development, ushering in a new era of clinical trials in which study techniques are as innovative as the drugs and devices they evaluate.

Parvin Fardipour, Ph.D., VP Statistical Methodology, ICON plc

With 27 years' clinical research experience Dr. Fardipour has significant drug development expertise with direct involvement in the clinical development of many compounds, and has successfully completed several NDAs in many therapeutic areas. Over the past 10 years she has worked in the adaptive design space to bring innovative approaches to clinical drug development and facilitate better and earlier decision making. She is an expert in designing, implementing and executing adaptive designs, enabling real-time learning and applying innovative designs across different therapeutic areas in both drugs and devices development.

Abdallah Ennaji, Senior Director Biostatistics, ICON plc

With a master's degree in statistics, Abdallah has over 20 years' experience in clinical research providing statistical expertise for the design and analysis of clinical trials in all major therapeutic areas, and has supported multiple development programs.

The authors can be contacted at: ***parvin.fardipour@iconplc.com; abdallah.ennaji@iconplc.com***