A sponsored whitepaper.

Strategies to Reduce Phase III Failure Rates

Alberto Grignolo
Corporate Vice President, Consulting
PAREXEL

Sy Pretorius
Chief Scientific Officer
PAREXEL

Half of all late-stage clinical trials fail due to lack of efficacy, safety or some combination of these two factors. Failure is a significant financial burden to biopharmaceutical companies, an immense frustration for investigators and a clinically substantial setback for patients who need new treatment options. And while it is impossible to prevent all Phase III failures, it is possible to adjust study concepts, designs and practices to reduce late-stage failure rates.

Theory suggests that Phase III failures should be few and far between. Early Phase studies are intended to be developmental, exploratory and, in principle, more likely to fail. As a result, the research and development pipeline that begins with as many as 10,000 compounds narrows to just five drug candidates entering clinical trials to produce a single agent approved by the Food and Drug Administration (FDA).

In the 1990s the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) defined Phase II studies as "exploratory" and Phase III studies as "confirmatory". This implies that Phase III studies should "confirm" the efficacy signals that presumably have been observed in Phase II. But two decades of data show that Phase II results are poorly predictive of Phase III results. In an independent analysis of oncology studies, for example, Phase III response rates tended to be lower than Phase II results for the same drug within the same chemotherapeutic regimen. The high expectations generated by Phase II success are too often crushed by Phase III failure.

Drug developers readily acknowledge this disparity. An audience poll during a recent webinar sponsored by global biopharmaceutical services organization PAREXEL International Corporation, found that 74 percent of attendees rated the risk of Phase III/late stage failure due to efficacy and/or safety problems to be a significant problem.

A 2014 FDA analysis of new molecular entities (NMEs) submitted for approval between 2000 and 2012 illustrates the problem. Of 302 NMEs submitted, 151 were approved in the first submission cycle. Of the 151 NMEs that failed the first cycle, 71 were approved following resubmission. The other 80 candidates were never approved.

Writing in JAMA, FDA authors noted that safety-only deficiencies accounted for 25.8 percent of first-cycle review failures and 35 percent of failed applications. Efficacy-only deficiencies accounted for 31.8 percent of first-cycle failures and 41.3 percent of failed applications. Even more troubling is the multiplicity of causes for efficacy failure.

Uncertainty about the optimal dose led to 15.9 percent of failures.

Faulty correlations between study populations and real-world populations accounted for 7.3 percent of failures.

SPONSORED BY

PAREXEL®

Applied Clinical Trials

A broad heading of "unsatisfactory endpoints" accounted for 13.2 percent of failures. Problems included lack of meaningful clinical benefit, premature time point, disagreement between FDA and principal investigators on the definition of successful treatment outcome, no improvement on overall survival and poor correlation of pathological endpoints with clinical benefits.

A recent PAREXEL analysis of 24 Phase III efficacy failures reported publicly by sponsors between mid-2012 and late 2014 found that about half of the failures involved oncology agents. Other therapeutic areas included Alzheimer's, diabetes, cardiovascular and autoimmune indications. These efficacy failures involved nearly 100,000 participating patients as well as time, effort and resources invested by sponsors, investigators, study sites and other study stakeholders.

Why Phase III Studies Fail

A September, 2014 conference at the European Center for Pharmaceutical Medicine identified a multiplicity of causes for Phase III trial failures which can be grouped in six broad categories: basic science, clinical study design, dose selection, data collection and analysis, operational execution and other causes. Any particular trial may be subject to multiple causes of failure.

Basic science weaknesses include animal models that are not entirely related to human disease, poor understanding of target disease biology, ineffective drugs or failures in translational medicine.

Clinical study design weaknesses include changes in patient definitions from Phase II to Phase III, insensitive outcomes measures or Phase II surrogate endpoints not confirmed by Phase III endpoints. Inappropriate study design can undermine the ability to show efficacy or the sample size may be too small.

Dose selection for Phase III may be inappropriate. If investigators or sponsors become too buoyed by Phase II results, they may rush to Phase III without fully exploring dose finding. Inadequate therapeutic indices may lead to suboptimal dosing.

Data collection and analysis problems include false positive signals from Phase II and overly optimistic assumptions about variability and treatment differences. There could be missing data, attrition bias, rater bias, errors in measurement methods or inappropriate statistical methods.

Operational execution could be marred by data integrity issues or GCP violations. Unexpected variations in recruitment or dropouts can affect results, as can protocol violations, missing data or unintentional unblinding.

The broad "other causes" category includes problems as diverse as futile studies of agents with no pathophysiologic benefit, a lack of critical examination of precedents or inadequate efficacy compared to standard of care.

Reducing the Risk of Phase III Failure

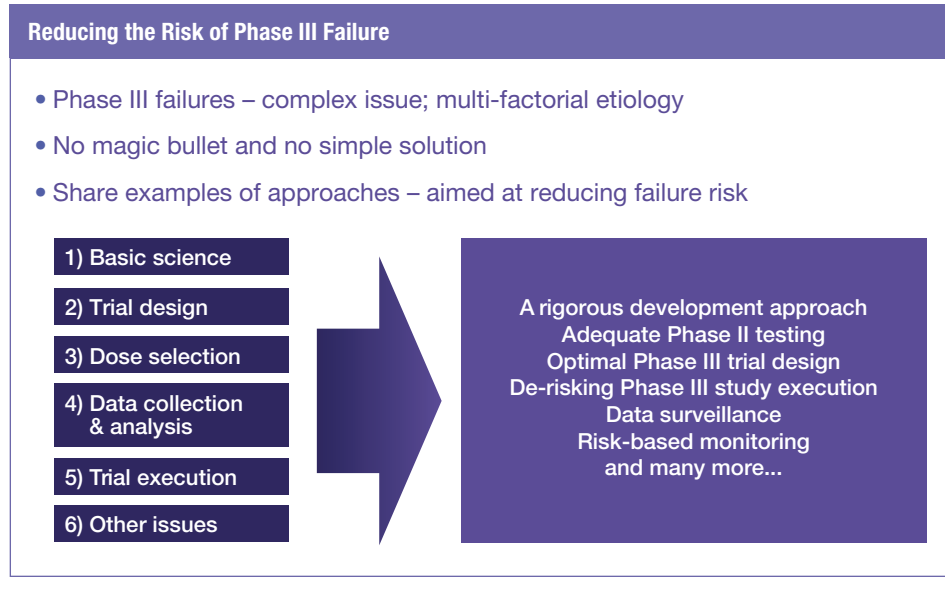
There are no simple fixes. Late-stage failures are a complex problem with multiple etiologies involving multiple combinations of external and internal factors. Several biopharmaceutical firms have developed strategies to help reduce the risk of failure. These strategies include a more rigorous development approach, adequate Phase II testing, optimizing Phase III study design, de-risking Phase III execution, data surveillance, risk-based monitoring and many more.

All of these approaches are designed to reduce the costly, frustrating, time-consuming and resource-devouring failures in Phase III. And while no one drug development team can be expected to apply all of these strategies, each strategy can help understand and address multiple causes of failure in late-stage clinical studies.

Rigorous Development

AstraZeneca applies its 5R Framework from the earliest stages of drug development. The five Rs help push researchers to develop the right target in the right tissue with the right safety profile for the right patients and the right commercial potential. The goal is to drive a higher degree of rigor and disciplined thinking throughout the development process.

There is also a sixth R, right culture. The Five Rs are most effective when organizational culture supports early and open admission of failure with evaluation of data in an open and transparent manner leading to appropriate decisions. It is



more time-effective, cost-effective, resource-effective and patient-effective to abandon an unpromising project at an early stage than to doggedly pursue failure to Phase III.

Adequate Phase II Testing

Pfizer developed a Three Pillar Framework to bolster Phase II testing. A rigorous application of cause and effect at Phase II can help identify failure early and reduce the risk of far more expensive failures at Phase III.

Pillar 1 demonstrates clearly and definitely the exposure of the drug candidate at the target site sufficient to elicit the intended pharmacologic effect over the intended time period.

Pillar 2 requires data clearly demonstrating that the drug binds to the specified target to produce the desired pharmacologic expression and modulation of the target site.

Pillar 3 is data clearly demonstrating that the pharmacologic modulation of the target site is functional, clinically relevant and expected.

Optimal Trial Design

PAREXEL and other drug development organizations have crafted multiple strategies to optimize trial design. Many step in the clinical trial design process can be enhanced, improved, strengthened and streamlined to reduce the risk of failure.

Study protocols typically consist of standardized components or building blocks: study objectives, population, proposed dose, study site procedures, statistical parameters, etc. To create alternative protocol scenarios, PAREXEL examines each of those building blocks from multiple perspectives, including regulatory, medical or therapeutic area, operations and statistics. Each alternative scenario is plotted by time and cost against a standard protocol. This structured analysis allows sponsors to evaluate the impact of different study designs to select the most appropriate combination.

Roche and other companies use modeling and in silico simulation to evaluate different trial designs and the potential

A Rigorous Development Approach: The Astrazenenca 5R Framework

Right target
- Strong link between target and disease
- Differentiated efficacy
- Available and predictive biomarkers

Right tissue
- Adequate bioavailability and tissue exposure
- Definition of PD biomarkers
- Clear understanding of preclinical and clinical PK/PD
- Understanding of drug-drug interactions

Right safety
- Differentiated and clear safety margins
- Understanding of secondary pharmacology risk
- Understanding of reactive metabolites, genotoxicity, drug-drug interactions
- Understanding of target liability

Right patients
- Identification of the most responsive patient population
- Definition of risk-benefit for given population

Right commercial Potential
- Differentiated value proposition versus standard of care
- Focus on market access, payer and provided
- Personalized health-care strategy, including diagnostics and biomarkers

Source: Drug Nature Reviews Volume 13, 419-431 16 May 2014

Adequate Phase Ii Testing: The Pfizer 3 Pillar Framework

- **Pillar 1:** Drug exposure at the target site of action is necessary to elicit a pharmacological effect over a desired time period.
- **Pillar 2:** Target occupancy is a prerequisite for expression of pharmacology and target modulation.
- **Pillar 3:** Functional modulation of the target is a prerequisite for expression of pharmacological activity to test the mechanism.

Source: Drug Discovery Today Volume 17, Numbers 9/10 May 2012

impact of the study disease, pharmacodynamics and pharmacokinetics, dropout rates and other elements that could affect the primary study objectives. Eli Lilly uses a mock physician practice or study site to simulate the operational conduct of a study. These mock ups, which are part of what Lilly calls Jam Sessions, are used to simulate and streamline procedures to make protocols more patient-centric and more likely to succeed.

None of these six approaches is foolproof, but all can work to increase the integrity of clinical studies and reduce the risk of failure.

Adaptive trial designs used by PAREXEL and others give sponsors the opportunity to revisit their initial assumptions at prespecified intervals using interim data. These early looks might allow a sponsor to drop a dosage arm that is ineffective or adjust sample sizes based on actual treatment effect differences. Biomarkers, patient enrichment strategies and assurance testing can also help optimize trial design to reduce the risk of failure.

De-risking Study Execution

The larger the study, the more likely that opportunities for deviations from protocol and lapses in data integrity will occur, increasing the risk of failure. Data-driven patient, country and site feasibility evaluation can help reduce the risks inherent in study execution.

Electronic health records, web listening, patient surveys and other data sources can all help study sponsors ensure that they are recruiting the right patients in the right countries and right sites to enhance the probability of success of a particular study. Countries or sites with low representation of patients with a specific health or physical characteristic might be poor choices for a specific study. Available data might reveal countries or investigators with a particularly poor, or stellar, track record for protocol compliance.

Ongoing data surveillance can help ensure that a study is on track and in compliance with protocol. Surveillance can also highlight deviations. Tracking patient recruitment can reveal localized imbalances between treatment groups indicating that a specific site is not enrolling the appropriate patients. This early warning provides an opportunity for intervention and correction during enrollment rather than discovering a fatal protocol breach after the database has been locked.

Risk-based Monitoring

All studies require monitoring, but monitoring based on risk can enhance study integrity. Safety experts at PAREXEL and other organizations have created multi-step algorithms to define critical data and processes, identify risks and risk thresholds, measure risks and align monitoring to mitigate risk and correct errors in order to track and document operational effectiveness.

None of these six approaches is foolproof, but all can work to increase the integrity of clinical studies and reduce the risk of failure. While Phase III failures can never be eliminated, thoughtful and deliberate assessment of potential points of failure in specific studies can help the industry evaluate and mitigate the risk of late-stage study failure.