Phase III Trial Failures: Costly, But Preventable

Alberto Grignolo, PhD, Sy Pretorius, MD

An examination of recent failures in Phase III studies and innovative approaches to reduce risk.

Thomas Henry Huxley (1825-1895) once stated, “The great tragedy of science—the slaying of a beautiful hypothesis by an ugly fact.” Indeed, in drug development not all ground-breaking scientific ideas translate into miracle cures. Biopharmaceutical companies around the globe strive to develop novel therapies to fulfill unmet medical needs, yet these efforts all too often die in development. Despite the numerous scientific, technological, and operational advances in R&D that would be expected to increase the efficiency and success of drug development, a significant number of clinical trials still fail to produce new, effective, and safe medicines.

Approximately 70% of Phase II trials are unsuccessful. Although this percentage might seem high, failure of early-phase trials is expected to some extent, as these trials are “exploratory,” “proof of mechanism,” and “proof of concept” trials in patients. What is unexpected, however, is the percentage of “confirmatory” Phase III trials that fail—about 50%. Theoretically, if early-phase trials provide the necessary criteria for moving a drug program to Phase III testing, relatively few Phase III trials should fail, but that is not the case.

In a recent PAREXEL analysis, we collected and evaluated data on 38 Phase III failures from mid-2012 through 2015 from a variety of publicly available sources (It is important to note that this list may not be exhaustive and may not include all of the Phase III trials that failed in this time period). As indicated in Table 1 (see facing page), these trials failed to meet primary or secondary efficacy endpoints; Phase III trials that failed due solely to safety issues are not included. These 38 failed trials collectively enrolled nearly 150,000 patients.

Peeling the onion: What are the drivers behind these Phase III failures?

Several groups have sought to analyze data from failed clinical trials to uncover the underlying trends and potential drivers of Phase III failures. Selected trends are discussed below.

1. A study by the Tufts Center for the Study of Drug Development (CSDD) evaluating clinical trials from 2000 to 2009 found that the three most common reasons that drugs or trials fail in Phase III of development are (see Figure 1 on facing page):
   - Efficacy (or rather lack thereof) — i.e., failure to meet the primary efficacy endpoint
   - Safety (or lack thereof) — i.e., unexpected adverse or serious adverse events
   - Commercial / financial — i.e., failure to demonstrate value compared to existing therapy

2. New molecular entities (NME) versus all drugs. A study by BioMedTracker (BMT) and the Biotechnology Industry Organization (BIO) evaluated R&D projects involving more than 9,500 different drug and biological products from 2004 to 2014. In each phase analyzed, the failure rates of trials involving NMEs were higher than trials involving all drugs, and the difference was greatest between Phase III development and submission of regulatory applications (See Figure 2 on page 38). Among the Phase III trials analyzed, only 61% of NME trials succeeded in moving to the application phase
The BMT/BIO study also found that clinical trial failure rates differ by therapeutic indication. Among the Phase III trials evaluated, oncology and cardiovascular trials had the highest failure rates (data not shown). Comparing failures for oncology and non-oncology trials specifically, failure rates for these indications differed more significantly at later stages of development (See Figure 3 on page 38). The oncology trials failed more often than non-oncology trials during the Phase II-to-Phase III transition and, in particular, from Phase III testing to regulatory submission (48% failure rate for oncology trials vs. 29% for non-oncology trials). The higher failure rate for oncology trials might be due to the inclusion of survival endpoints and the need to show efficacy by an improvement in overall survival.

4. Failure rates differ by type of molecule. A study by the Tufts CSDD found that the probability of success for clinical trials of small molecules is lower than for trials of large molecules (see Figure 4 on page 39). Among the Phase III trials analyzed, only 61% of studies involving small molecules progressed from Phase III testing to the regulatory application phase (39% failure rate). Trials of large molecules, however, were more successful in moving to the application phase, with a success rate of 74% (26% failure rate). The study also found that within the large-molecule subtypes, failure rates for recombinant proteins and monoclonal antibodies were similar overall, but recombinant proteins failed more often than monoclonal antibodies in the transition from Phase III testing.
to regulatory review (34% failure rate for recombinant proteins vs. 13% for monoclonal antibodies; data not shown).

The European Center for Pharmaceutical Medicine (ECPM) organized a seminar in September 2014 titled “Why Clinical Trials Fail?” Based on presentations delivered by various experts at the seminar, we have grouped the drivers behind Phase III failures identified by these experts into six categories: basic science, clinical study design, dose selection, data collection and analysis, operational execution, and other. Table 2 (see facing page) provides examples of specific events in each category that could lead to Phase III failure.

Impact of Phase III failures: Human and financial
The effects of Phase III failures are significant, with the most poignant being the impact on human lives and financial cost. First and foremost is the negative impact on the thousands of patients who enroll in clinical studies hoping to find a viable treatment option. As mentioned previously, PAREXEL’s analysis revealed that in a relatively narrow time frame (2012-2015), nearly 150,000 patients had enrolled in Phase III trials that eventually failed. Most of these patients had cardiopulmonary diseases, diabetes mellitus, or cancer. Patients who participate in clinical trials often have difficult-to-treat or later-stage diseases with few, if any, standard treatment options. For many of these patients, time may be running out, and a clinical trial is the last option for a potential cure or prolongation of life. Thus, the impact of failed drug trials can be widespread and may worsen patient prognosis, quality of life, and emotional well-being.

Not surprisingly, failures of large Phase III trials also result in a material financial burden to biopharmaceutical companies and are a contributor to poor capital productivity. Considering the substantial financial investment in R&D, few new drugs are being approved by regulators. One study found that for every billion U.S. dollars spent on R&D, the number of new drugs approved has decreased by approximately 50% every nine years since 1950 (inflation-adjusted). And, the financial consequences of failed drug development often go beyond R&D. An examination of the most impactful Phase III failures of 2014 found that these failures resulted in the termination of drug programs and employees as well as financial losses for investors, as shown in Table 3 (see page 40; note that only some of these failures are also listed in Table 1). With slower growth in R&D spending, it is important that biopharmaceutical companies use the limited funds and resources efficiently and effectively to get drugs to market and provide new medicines that meet unmet medical needs.

Selected strategies to reduce the risk of late-stage failure
Although no simple solution exists yet for preventing Phase III failures to confirm the efficacy of a new drug or biologic, several approaches have been and are currently being developed to reduce the risk of such failure. Given our unique perspective in working with hundreds of companies across thousands of clinical trials and compounds, we and numerous colleagues at PAREXEL are exposed to these approaches on a daily basis and briefly highlight some of these in the next sec-

---

**Figure 2.** The difference in rate success of trials involving NMEs was greatest between Phase III development and submission of regulatory applications.

**Figure 3.** Failure rates for oncology and non-oncology trials differed more significantly at later stages of development.
TRIAL DESIGN

These strategies can be implemented during the entire development process, in specific phases of development, and/or during clinical trial design.

Applying more rigor to the overall development process
Over the past few years, most drug development companies have established and adopted more disciplined protocol, progress and portfolio review frameworks. In 2011, AstraZeneca aimed to overhaul its R&D process to improve the health of the organization and increase the chance of success of its Phase III trials. By evaluating its small-molecule drug projects over a five-year period (2005–2010), AstraZeneca identified the factors associated with project success and developed a framework that now drives its development process. The 5R framework guides R&D teams in identifying the right target, the right tissue, the right safety, the right patients, and the right commercial potential (See Figure 5 on page 41).

Having the right culture—i.e., a culture where the facts and data are confronted with brutal honesty, and courageous kill-early decisions are encouraged where appropriate—is an important, additional dimension to this framework. Applying more rigor and discipline to the development process holds potential for weeding out likely failures earlier in the process, thereby reducing Phase III failure rates.

Adequate Phase II testing
Many Phase III trials fail because of a fundamental lack of understanding of the mechanisms of action of NMEs. To address this issue and improve the chance of NMEs being successful during Phase II testing, Pfizer performed an analysis of 44 Phase II programs during a four-year period (2005-2009) to identify factors associated with success. The result of that analysis led to the “Three Pillars of Survival” framework that helps the R&D teams determine three key elements that increase the likelihood of an NME surviving Phase II testing and moving on to Phase III testing: Pillar 1 involves exposure of the drug to the target site of action, Pillar 2 involves binding of the drug to the target, and Pillar 3 involves expression of pharmacology (see Figure 6 on page 42). Rushing to get to Phase III without adequate understanding of these three criti-

### Failure Triggers

<table>
<thead>
<tr>
<th>DRIVERS OF FAILURE</th>
<th>EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inadequate basic science</td>
<td>• Beneficial effects in animal models not reproduced in humans</td>
</tr>
<tr>
<td></td>
<td>• Poor understanding of target disease biology</td>
</tr>
<tr>
<td>Flawed study design</td>
<td>• Patient population definition changed from Phase II to Phase III</td>
</tr>
<tr>
<td></td>
<td>• Phase II surrogate endpoint not confirmed by Phase III clinical outcomes</td>
</tr>
<tr>
<td></td>
<td>• Insufficient sample size</td>
</tr>
<tr>
<td>Suboptimal dose selection</td>
<td>• Inadequate dose finding in Phase II</td>
</tr>
<tr>
<td></td>
<td>• Poor therapeutic indices</td>
</tr>
<tr>
<td>Flawed data collection and analysis</td>
<td>• Phase II “false positive” effects were not replicated in Phase III</td>
</tr>
<tr>
<td></td>
<td>• Overoptimistic assumptions on variability and treatment difference</td>
</tr>
<tr>
<td></td>
<td>• Missing data; attrition bias; rater bias</td>
</tr>
<tr>
<td></td>
<td>• Wrong statistical tests; other statistical issues</td>
</tr>
<tr>
<td>Problems with study operations</td>
<td>• Data integrity issues; GCP violations</td>
</tr>
<tr>
<td></td>
<td>• Recruitment, dropouts, noncompliance with protocol</td>
</tr>
<tr>
<td></td>
<td>• Missing data; unintentional unblinding</td>
</tr>
<tr>
<td>Other</td>
<td>• Insufficient landscape assessment of current standard of care and precedents</td>
</tr>
</tbody>
</table>

Source: European Center for Pharmaceutical Medicine; PAREXEL Analysis

**Table 2.** Specific events in key areas of clinical trial conduct that can result in Phase III failure.

### Success Rates by Molecule Type

- **Transition probability**
  - Phase I-II: Small molecule = 63%, Large molecule = 84%
  - Phase II-III: Small molecule = 38%, Large molecule = 53%
  - Phase III-NDA/BLA Sub: Small molecule = 61%, Large molecule = 74%
  - Phase NDA/BLA Sub-NDA/BLA App: Small molecule = 91%, Large molecule = 96%
  - Phase III-NDA/BLA App: Small molecule = 13%, Large molecule = 32%

Source: Tufts Center for the Study of Drug Development (CSDD)

**Figure 4.** According to a study, the probability of success for clinical trials of small molecules is lower than for trials of large molecules.
Optimized Phase III Clinical Trial Design

Flaws in clinical trial design are a major driver of Phase III failures. Several strategies have been developed for optimizing trial design. The next section aims to highlight a few of them:

- **More rigorous protocol review and optimization** is an approach that many companies are employing in varied degrees ranging from stage gate reviews and sign-offs by various internal committees to live, in-practice simulations of protocols to identify and resolve potential glitches proactively. Typically, a number of trade-off decisions need to be made in the compilation of most Phase III protocols—e.g., will the protocol include a specific secondary objective that has commercial value but potentially prolongs the duration of the study? Identifying and quantifying the impact of these trade-offs is helpful in designing better protocols. Likewise, in the right capable hands, the vast amount of data available in public sources (e.g., Clinicaltrials.gov) can often be harnessed to determine what worked well and what did not.

- **Modeling and simulation** have contributed to regulatory approval and labeling decisions in recent years and are currently being employed as a potential solution for mitigating late-stage failure risk. In a review of 198 applications submitted to the FDA between 2000 and 2008, modeling and simulation was found to provide pivotal or supportive insights into effectiveness and safety that contributed to the approval of 126 applications. Additionally, it provided information on dosage, administration, and safety that supported 133 labeling decisions. The FDA encourages sponsors to incorporate quantitative modeling and trial simulation in the drug development plan and to seek regulatory guidance on these strategies through participation in an end-of-phase 2A (EOP2A) meeting. The EOP2A meeting provides the sponsor with an opportunity to discuss modeling and simulation plans and receive feedback from the FDA on how to optimally implement these strategies to quantify the exposure-response relationships and select appropriate doses for Phase III trials. Modelling and simulation is currently being used more broadly than just for the selection of an optimal dose. Modeling and simulation of various clinical trial designs is one such example that is being used in an effort to design optimal Phase III studies.

### Financial Consequences of Phase III Failures

<table>
<thead>
<tr>
<th>SPONSOR</th>
<th>DRUG</th>
<th>INDICATION</th>
<th>CONSEQUENCES OF FAILURE</th>
</tr>
</thead>
</table>
| Exelixis | cabozantinib | metastatic castration-resistant prostate cancer | • 150 jobs (65% of workforce) terminated  
• $6-$7M in termination costs (95% charged in 3Q 2014; 5% in 1Q 2015)  
• Reversed $2.1M of previously recorded stock-based compensation expenses  
• Cancelled 692,896 stock options due to unrealizable performance objectives |
| OncoGenex Pharmaceutical Industries | custirsen | metastatic castration-resistant prostate cancer | • Share price fell 60%  
• Collaboration with Teva terminated |
| GlaxoSmithKline | MAGE-A3 | Non-small cell lung cancer | • Share price fell almost 2%  
• Development halted |
| Nymox Pharmaceutical | NX-1207 | benign prostatic hyperplasia | • Share price fell nearly 82%  
• Development halted |
| Regado Biosciences | pegnivacogin + anivamersen | coronary artery disease undergoing percutaneous coronary intervention | • Share price fell 60%  
• 20 jobs (60% of workforce) terminated  
• $2M in termination costs charged in 4Q 2014 |
| Eli Lilly | tabalumab | systemic lupus erythematosus | • Development halted  
• $63M in termination costs charged in 3Q 2014 |
| Merck & Co./Endocyte | vintafolide | platinum-resistant ovarian cancer | • Share price fell 62%  
• Withdrawal of marketing applications for vintafolide & companion imaging products |


**Table 3.** The financial and other company-related losses resulting from a selection of Phase III trial failures in 2014.
Trial Design

THE ‘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 PO 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 target liability
- Understanding of secondary pharmacology risk

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 provider
- Personalized healthcare strategy, including diagnostic and biomarkers

Source: Cook et al.

Figure 5. AstraZeneca’s framework to help guide R&D teams, focusing on the five “R’s” of the development process.

- Adaptive designs are gaining momentum as a way to reduce failure risk. This is, in large part, due to the fact that these adaptive designs provide an opportunity to assess interim data and to sense-check some of the initial uncertainties or assumptions that were made at the outset of the trial.15,16 More importantly, these provide an opportunity to change course during a trial and to correct these incorrect assumptions in a prospectively planned manner that does not jeopardize statistical validity and the operational integrity of the trial.17

- Biomarkers are being used increasingly to assess efficacy in a rapid and more objective manner.14 A particular challenge associated with the use of biomarkers in clinical trials is the requirement to validate these biomarkers as relevant disease-modifying endpoints and correlate changes in these endpoints to clinically significant changes in disease progression.18 Likewise, enrichment strategies often involving genotyping are being used increasingly to optimize the proposed study population to those individuals most likely to respond to the treatment under investigation.20

De-risking study execution

A number of approaches are currently being used to mitigate study execution risks. These include:

- Leveraging data from a variety of sources (e.g., electronic health records [EHR], prior site performance, central labs, etc.) to ensure that the selection of patients, sites, and countries are optimal. The impact of EHRs on clinical trials is carefully tracked by numerous interested parties, and it is not difficult to imagine the potentially transformative impact that these might have on clinical trials. At the same time, many industry insiders hold a somewhat skeptical view—especially about the accuracy of data in these systems.

  - Ongoing surveillance of the quality of data being collected during a Phase III trial and tying this to a properly planned risk-based monitoring protocol allows companies to set certain quality triggers that would result in increased monitoring when and where needed. This is helpful in reducing execution risk, as errors at the site level are often discovered too late in the process—i.e., once the study is clinically completed and the database is locked. Often, at this time, very little can be done to salvage the situation and improve the quality and potential for success in the trial.

Completeness and clarity of submissions and interaction with regulatory agencies

Many drugs fail to get approved because the information submitted to the regulatory agency is insufficient to allow for a determination on safety and efficacy, not because the drug is actually unsafe or ineffective. A review of marketing applications for NMEs submitted to FDA from 2000 to 2012 revealed that about 50% of those reviewed failed to obtain approval during the first-cycle.2 However, nearly 50% of these failed applications were eventually approved on resubmission after applicants addressed the FDA’s concerns related to safety, manufacturing, and labeling. Delayed approvals and non-approvals may be reduced or prevented by seeking advice from regulators during the drug development process. The FDA offers sponsors and applicants the opportunity to schedule formal meetings at critical points in the development and regulatory process to discuss the plan and identify areas of concern that may need adjustment.21

Other bold/promising/theoretical/philosophical strategies that could reduce Phase III failures

Other areas that are currently being explored and could hold potential in reducing Phase III failure risk include:

- Replacement of the current gold standard, the randomized controlled trial, with real-world evidence
- Wearable devices that collect real-time data22
- Adaptive licensing
- Next-generation sequencing and improved understanding of the genetic basis of disease
- Basket/master protocols

Although Phase III failures cannot be eliminated, the risk of these failures can be reduced. The strategies discussed in this article may provide sponsors with some of the tools aimed at minimizing the risk of failure in their drug development plans.

Conclusion

We believe that the current failure rate in Phase III studies is unacceptably high, and that industry is keen on reducing this failure risk, although some in industry may believe that fail-

The ‘Three Pillars of Survival’

- 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: Morgan et al.

Figure 6. Pfizer’s framework to help determine three key elements that raise the likelihood of an NME surviving Phase II testing and moving on to Phase III.

ure is the price to be paid for innovation. As a first step, it is important to understand the reasons and root causes driving these failures. Our research identified recently failed Phase III studies that have enrolled nearly 150,000 patients. Based on data from our analysis and others, we have listed the main reasons why Phase III trials fail. In addition, and given our unique perspective in working with hundreds of sponsors across thousands of trials, we have highlighted some of the approaches that pharmaceutical companies are implementing in an effort to reduce these costly late-stage failures. Along with our colleagues in the pharmaceutical industry, we are optimistic about the potential of some or all of these approaches to improve the Phase III success rate.

Acknowledgements
Dawn Giles from PAREXEL International and Jennifer Swanson from JS Medical Communications, LLC, contributed to the development of this article by providing assistance with research, writing, and editing.

References

Alberto Grignolo, PhD, is Corporate VP, Global Strategy, PAREXEL; Sy Pretorius, MD, is Chief Scientific Officer, PAREXEL.