Randomized controlled trials (RCTs) represent the gold standard for testing a clinical hypothesis. Since most regulators demand hard evidence in support of new drug applications, and since most pharmaceutical companies try to follow the lowest risk path to product approval, it is not surprising that most Phase II and III studies follow this design. Although most sponsors favor RCT designs for their registration trials, they frequently choose other approaches for their Phase IIIB and IV studies to save time, minimize cost and avoid ethical challenges associated with conducting experimental treatment protocols in patients. Choosing an appropriate design often means balancing the need for rigorous
scientific proof against a need for timely information, or against the morality of assigning patients to receive treatments of significantly different risk or efficacy. For example, it would be unethical to measure the effect of a new treatment on healthcare utilization by asking one group to take the new product and another group to use the incumbent market leader which has already been proven clinically inferior in earlier stage studies.

This chapter briefly describes different study designs commonly used in Phase IIIB and IV research. It begins with a discussion of randomized controlled trials and different options for minimizing reporting bias, and then turns to a discussion of non-experimental approaches to clinical research using retrospective, cross-sectional and prospective observational study designs. Those wanting more details are encouraged to seek further information from a specialist text on clinical trials design, biostatistics, or epidemiology.1

Randomized Controlled Studies

Patients participating in RCTs are randomly assigned to one of several treatment groups. This process of “randomization”, first proposed by Sir Ronald A. Fisher in 1919, allows researchers to establish a control experiment by ensuring that all variables except for the treatment received are kept roughly constant between patient groups. Doing so increases the chance that different outcomes observed between the groups can be attributed to the different treatments that they received.

To minimize the effect of measurement and reporting bias on study results, some studies require that neither the patient nor the investigator know which treatment the patient is receiving. This process, known as double-blinding, increases the statistical strength of a study and is the most highly regarded method for conducting randomized clinical research but can be unpopular with research naive physicians and patients who are not comfortable with losing control over treatment decisions. In addition, sponsors can face significant delays and costs associated with manufacturing treatment formulations that must be indistinguishable from each other for double-blinding to work. Sometimes researchers compromise by allowing the physician (but not the patient) to know what treatment the patient is receiving, a process known as single-blinding. Single-blinding is often used to evaluate surgical interventions or drugs that need frequent dose titration since it is not possible for the investigator to avoid knowing what treatments are being used.

In contrast, open-label studies allow both patients and physicians to know which

The PROBE Design

The prospective, randomized, open, blinded endpoints (PROBE) design has become popular for research involving large numbers of patients or comparison of treatments with very different routes of administration, both of which make double-blind trials very complex and costly. PROBE designs have been well established in cardiovascular research, most particularly in hypertension studies, following the success of the ASCOT trial, which randomized over 19,000 patients to one of two different open-label antihypertensive treatment programs. The open-label design was necessary in this case for three reasons: because physicians were required to tightly control blood pressure through frequent dose titration (hard to do if the physician does not know which drugs they have to titrate), because physicians were required to closely monitor for electrolyte disturbances in patients receiving diuretic treatment, and (3) because the large study size required that study designers focus on reducing workload burden and cost.

Subsequent meta-analyses of both PROBE and double-blind placebo-controlled hypertension studies have demonstrated the statistical equivalence of their results, and provided validity of the PROBE design as a powerful research alternative.3
treatment each patient is receiving – introducing the possibility of reporting bias but minimizing the cost and time-line for obtaining study results and expanding the number of physicians that might be interested in participating in the research. Patients and their physicians usually feel more comfortable knowing what drugs are being taken, and it is easier for sponsors to use their normal supply of drug in the trial.

After the clinical efficacy of a new medicine has been proven in Phase II and III studies, many physicians want to know about the net effectiveness and safety of the product – including any effect that its packaging, formulation, and brand placebo effect may have on its tolerability, usability, and compliance. Understanding that a treatment is more than just a molecule, many physicians value health outcomes demonstrated through open-label research because it attempts to characterize the value of a product when used as it would be in the real world. Wherever possible, open-label randomized studies attempt to avoid reporting bias by using objective machine measurable endpoints, such as total cholesterol, temperature, body mass index, or blood pressure. To avoid the possibility of test results feeding back to influence treatment, many researchers choose to keep patients and their physicians blind to the results of their individual endpoint measures – creating what is popularly referred to as the PROBE (prospective, randomized, observational blinded endpoints) design.

Observational Methodologies

Randomized studies involve the setup of an experiment to compare treatment outcomes between different patient groups with patient treatment dictated by a study protocol rather than a physician’s standard of care. Although an RCT represents the gold standard for testing a clinical hypothesis, there are many different non-experimental study designs that can still yield valuable clinical information much more quickly, straightforwardly, cheaply – and sometimes more ethically – than this experimental approach.

These non-experimental designs, also known collectively as observational studies, can be retrospective (using data that has already been captured), cross-sectional (capturing patient data at one point in time), and prospective (enrolling and tracking patients over time). Following is a brief discussion of each approach, with case examples to illustrate the role of these different study designs in the clinical research process.

Retrospective Studies

Retrospective studies use historical data to populate a study database; the most common example, the chart review, captures data from patient files. Retrospective studies are much cheaper than prospective or cross-sectional studies, which typically involve the prospective enrollment and consent of patients. Many retrospective studies simply require that sites transpose de-identified summary data from their existing files into a study database for analysis, in most cases avoiding lengthy administrative processes. Retrospective studies can be completed even more quickly if researchers choose to use claims databases or existing clinical registries to support ongoing data analyses.

Because retrospective studies usually involve different data points collected at multiple points in time for each patient, they are recognized as being more effective at inferring causality than a cross-sectional study but are frequently challenged by critics who can only be satisfied by RCT data. On August 25, 2004, Dr. David Graham presented the results of an FDA-funded retrospective analysis of Kaiser Permanente health record claims. The highly publicized study suggested that patients receiving Merck’s Vioxx treatment had experienced a significantly greater rate of cardiovascular events than those receiving older non-steroidal agents. Not surprisingly, even though the media reaction to this data was intense, many
industry analysts and key opinion leaders initially responded cautiously to this study because Graham’s analysis was not an RCT.

Graham’s announcement of his retrospective Kaiser Permanente data analysis provides an excellent example of how these studies can establish interesting hypotheses to direct further research; less than 14 days later, following an analysis of data from their own randomized Vioxx studies, Merck decided the link between Vioxx and cardiovascular risk had sufficient scientific validity and announced the biggest global product recall in history.

**Case-Control Case Study**

To identify risk factors for intrauterine fetal death (IUFD), researchers from Nottingham City Hospital compared data from 161 singleton stillbirth pregnancies that occurred between 1991 and 1997 with data from 499 randomly selected live births that occurred during the same period. The study identified several factors as being positively associated with IUFD – small size for gestational age, maternal body mass index, maternal age, and maternal type O blood group – but revealed no association between stillbirth rate and maternal ethnicity, Rhesus status, fetal sex, or smoking.

**Case-control design.** The case-control design involves identifying patients with a given health outcome ("cases") and those without the outcome ("controls"), and then looking back in time to compare the frequency of an exposure in the case group to the control group. The case-control design is commonly used to study rare health outcomes and is sometimes the only ethical way to investigate an association.

**Ecological design.** An alternative retrospective approach, population-based rather than individual-subject based, involves aggregating patient data at investigational sites and then transmitting only a summary of this data back to the study center for analysis. In most regions, the process of aggregation at a group level avoids privacy issues. For example, an investigator may wish to study the mortality rate associated with different types of chemotherapy treatment. Rather than asking five oncology centers to submit patient-level data, the investigator may simply ask each center to provide high-level data regarding the number of patients being treated with different chemotherapy programs and the mortality rate associated with each group.

While retrospective studies are relatively quick and cheap to implement, the quality and completeness of data collection is dependent upon the quality and appropriateness of data recorded in the medical history or other data repositories. It is often impossible to go back and query data, and even then data points are limited to those recorded during the standard consultation process. Retrospective research concludes with many values still missing and data queries unresolved.

**Cross-Sectional Studies**

Cross-sectional studies, as their name suggests, involve collection of data from a patient at a single point in time. Usually, cross-sectional studies try to enroll patients over a short period of time to prevent time lapse over the data collection period from influencing results. For instance, an allergy study will yield lower severity scores if patient enrollment spans beyond spring.

Cross-sectional studies represent a cheaper and quicker way to obtain data than prospective research while offering more control over question design and data completeness than a retrospective approach. Cross-sectional study designs are most commonly used for determining prevalence (the number of cases in a population at a given point in time), but they can be used to identify possible associations without a long trial. Sometimes the inference of causality by a cross-sectional study is sufficient to change a clinical behavior. For example, the mechanism of action for a new painkiller may suggest a theoretical
Cross-Sectional Case Study

Interested in understanding more about the functional disability of rheumatoid arthritis, researchers from Norway conducted a cross-sectional study of 706 European patients who had had rheumatoid arthritis for at least four years. They assessed the functional disability of each patient at a point in time and collected information regarding possible risk factors and markers for disease activity. The study demonstrated that female sex, high erythrocyte sedimentation rate (ESR), and disease duration strongly and independently correlated with functional disability, while the presence of rheumatoid factor, joint damage as observed on X-ray, increasing age, and education did not. The results of this study have helped physicians recognize that women are at particularly high risk of disability even in the early stages of the disease, which has justified more aggressive approaches to early treatment.

Naturalistic Case Study

Sometimes the effectiveness of a treatment is dependent upon a much wider range of factors than molecular efficacy alone. This is particularly so in the case of dementia, where the quality of non-medical patient care can have a significant effect on patient health outcomes and the burden the disease places on families. To assess the impact of Reminyl (galantamine) on behavioral disturbances and associated caregiver burden when used under naturalistic conditions, investigators from Switzerland tracked the outcomes of 124 patients initiated on the treatment. The study demonstrated significantly reduced behavioral disturbances in this population after three months and a measurable reduction in caregiver burden. For primary care physicians who are managing the health and stress of caregivers as well as patients, this data is likely to be more influential as a guide to treatment than RCT results supporting a small clinical effect on cognitive function.

Prospective Observational Studies

A prospective observational study, also known as a prospective cohort study, is one that simply monitors a group (or cohort) of patients over time without mandating any treatment intervention. Most industry references to “observational research” usually relate to this prospective approach, where the decision to treat is the physician’s alone and any treatment prescribed during the study would have been considered by the physician whether or not the patient was a study participant. Observational studies can involve a schedule of follow-up visits and tests that are not part of usual clinical practice, and sometimes involve blinding the patient but obviously not their treating physician, to the treatment they receive.

These studies – particularly for approved products – present little risk to patients since they allow the treating physician to remain in control of all treatment decisions. As a result, they require less monitoring activity (and therefore less cost) than an RCT to ensure data quality and patient safety, and patient recruitment is much easier. Prospective observational studies often represent a compromise for sponsors who require very large amounts of high quality targeted data not available through other sources, but who cannot justify the cost or timeline associated with gastrointestinal benefit. A cross-sectional study of patients receiving painkillers may reveal that gastric pain is less prevalent in patients receiving this new drug than in those receiving other drugs on the market. While this is clearly not proof of its relative safety, these results may be enough to make many prescribers preferentially prescribe the new drug to patients with gastrointestinal irritability.
with conducting an RCT.

**Naturalistic Designs**

A naturalistic (or actual use) study is a special case of a prospective observational study that not only avoids mandating any treatment intervention but also avoids scheduling non-standard visits, examinations, or investigations that are not otherwise performed in routine practice. People often mistakenly use the term “naturalistic” to describe studies that are actually observational. A naturalistic study, by definition, is always observational; an observational study is not always naturalistic.

Sponsors perform naturalistic studies when it is important that the conduct of a study itself should not influence patient care patterns and treatment outcomes. A study that offers free study drug is not naturalistic, since patients may be more compliant and less likely to discontinue treatment than if they purchased drug through a formulary. A study that involves more visits than those normally associated with routine patient care is not naturalistic, since the increased frequency of review may change the way a physician treats patients. Similarly, a study that mandates particular tests is not naturalistic, since the results of tests performed outside naturalistic practice may influence patient care. Would a physician ignore a high cholesterol result obtained through a clinical trial even if the patient were not otherwise scheduled for a cholesterol test for another six months?

**Parallel-Cohort Designs**

Whereas most prospective observational studies track clinical outcomes for a single treatment group, it is occasionally useful to follow multiple treatment groups, or cohorts, in parallel. By comparing outcomes from two different groups, researchers can use an observational approach to answer complex clinical questions.

Using tight inclusion/exclusion criteria to ensure that confounding characteristics are equally balanced between study groups, researchers can establish a prospective observational study that approximates a randomized design. For example, an observational protocol may ask a physician to enroll the first three healthy males, aged 30–35, receiving drug X, and the first three healthy males with the same clinical characteristics taking drug Y. As long as both populations have similar characteristics at the time of analysis (homogeneity generally improves as study samples increase), comparative analysis through direct comparison or case matching may be possible. Researchers can apply a number of statistical techniques to maximize the significance of conclusions drawn.

**Hybrid Studies**

Many strategic research studies patch together multiple design approaches, attempting to capitalize on the speed, cost-efficiency, and empirical strength of different designs by combining them within a research program. Such approaches include mixing cross-sectional designs with retrospective data capture, thereby approximating a prospective observational approach without the time delay. Other approaches include running a simple prospective study across a large number of sites in parallel with a set of smaller, more complex randomized controlled sub-studies at a subset of sites. Mixing designs in this manner can be cost effective and powerful but can also introduce risk to the analytical integrity of the project if not conducted by individuals with experience in study design.
Hybrid Case Study

Establishing the effectiveness of a treatment in reducing patient mortality or morbidity can be time consuming when the therapy is used to prevent illness rather than to treat disease. Sometimes it is necessary to follow patients for five to ten years before the real impact of a treatment can be established. One company, keen to rapidly obtain this data in support of their preventative cardiovascular treatment, commenced a hybrid study including both retrospective and prospective components. Primary care physicians were asked to identify a pool of patients who had already received at least three years of treatment with the active drug, as well as a much larger secondary pool of patients who had received treatment with a competitive product during the same period. A case matching process was undertaken based on patient characteristics at the time that therapy was originally initiated. Prospective follow-up of clinical markers and patient outcomes is now taking place over a two-year period, with regular comparative analysis to detect deviation in survival outcomes. By taking this approach, the sponsor is likely to generate strong data in support of their product years before results from their long-term RCT are available.

Conclusion

The relative benefits and costs of different clinical trial designs available to the researcher in Phase IIIB and IV are summarized in the tables below, including their capacity to answer a range of questions (breadth).

<table>
<thead>
<tr>
<th>Type of Trial</th>
<th>Time</th>
<th>Cost</th>
<th>Breadth</th>
</tr>
</thead>
<tbody>
<tr>
<td>QUICKLY (But less robust)</td>
<td>Cross-Sectional</td>
<td>2-6m</td>
<td>$0.5-2m</td>
</tr>
<tr>
<td>ROBUSTLY (But less quickly)</td>
<td>Prospective Observational Multi-Arm</td>
<td>12-18m</td>
<td>$1-5m</td>
</tr>
<tr>
<td>COMPROMISE</td>
<td>Prospective Observational Single-Arm</td>
<td>6-18m</td>
<td>$0.5-5m</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of Trial</th>
<th>Time</th>
<th>Cost</th>
<th>Breadth</th>
</tr>
</thead>
<tbody>
<tr>
<td>WITH CONCLUSIVE PROOF</td>
<td>Randomized Control (RCT)</td>
<td>1-3 yrs</td>
<td>$1-20m</td>
</tr>
<tr>
<td>WITH STRONG PROOF</td>
<td>Retrospective Case Control</td>
<td>0-6m</td>
<td>$0.2-1m</td>
</tr>
<tr>
<td>... BUT DON'T HAVE THE DATASET</td>
<td>Prospective Observational</td>
<td>6-18m</td>
<td>$0.5-5m</td>
</tr>
<tr>
<td>... AND DON'T HAVE THE TIME</td>
<td>Cross-Sectional</td>
<td>2-6m</td>
<td>$0.5-2m</td>
</tr>
</tbody>
</table>
Strategic research can benefit from the wide variety of design options available to investigators. While Phase IIIB and IV researchers frequently use the most rigorous randomized, controlled, double-blind designs, they also commonly use observational studies and open-label designs. These are often more compatible with normal clinical workflow and may provide sufficiently relevant data more quickly, easily, and cheaply than would be possible taking a randomized double-blind controlled trial approach.

References


* Note: All citations to websites listed in this book were verified in May of 2005.