



ARIEL D. STERN

SARAH MEHTA

Adaptive Platform Trials: The Clinical Trial of the Future?

In July 2017, Dr. Brian M. Alexander, disease center leader for radiation oncology at Boston-based Dana-Farber Cancer Institute's Center for Neuro-Oncology, was preparing to launch a new type of clinical trial—an *adaptive platform trial*—to study potential therapies for glioblastoma (GBM), an aggressive form of brain cancer. GBM, which had recently made headlines when U.S. Senator John McCain was diagnosed,¹ was among the deadliest of cancers and had no known cure. Most people with GBM died within 15 months of diagnosis,² and just 15% survived past five years.³

Alexander had grown frustrated with the limited treatment options for his GBM patients. He believed that the standard way in which researchers tested the effectiveness of new treatments—traditional randomized controlled trials (RCTs)—was limited in many ways. RCTs randomly assigned patients to either a “treatment” or “control” arm: patients in the treatment arm received the new therapy while those in the control arm received either the standard existing treatment or a placebo. Researchers then compared health outcomes between the two arms over time to determine whether the new therapy was effective and to identify side effects. While statistically rigorous and still considered the “gold standard” in clinical research, traditional RCTs were time-consuming, costly, and limited to testing just one drug at a time.

An overwhelming minority of patients with GBM were enrolled in clinical trials, creating a frustrating dilemma for both patients and their oncologists. As Alexander lamented, “I am disheartened each time a patient is unable to enroll in a trial. Outside of a clinical trial, we have no systematic way to learn from a patient’s journey and, even worse, they have no opportunity to try different, potentially better, therapies than the meager standard of care we have to offer. There must be a better way.”

To that end, for the past three years, Alexander had been working closely with a group of like-minded oncologists and statisticians to design an adaptive platform trial called Adaptive Global Innovative Learning Environment for Glioblastoma (GBM AGILE) in the hopes of identifying effective therapies more quickly. Unlike traditional RCTs, adaptive platform trials maintained several “treatment arms” to simultaneously and dynamically study the effects of multiple unique drugs for a given disease. As such, they had the potential to fundamentally change the clinical research process,

Professor Ariel D. Stern and Case Researcher Sarah Mehta (Case Research & Writing Group) prepared this case with the assistance of Richard G. Hamermesh who provided significant support and feedback, and Katherine Lofgren who provided statistical research assistance for the case appendix. It was reviewed and approved before publication by a company designate. Funding for the development of this case was provided by Harvard Business School and not by the company. HBS cases are developed solely as the basis for class discussion. Cases are not intended to serve as endorsements, sources of primary data, or illustrations of effective or ineffective management.

Copyright © 2017, 2018 President and Fellows of Harvard College. To order copies or request permission to reproduce materials, call 1-800-545-7685, write Harvard Business School Publishing, Boston, MA 02163, or go to www.hbsp.harvard.edu. This publication may not be digitized, photocopied, or otherwise reproduced, posted, or transmitted, without the permission of Harvard Business School.

making clinical trials for new cancer drugs more efficient, more available, and more ambitious in scope. They also used statistical techniques to assign a higher proportion of patients to the treatment groups from which they were more likely to benefit. In one study comparing clinical trial types, for example, researchers estimated that an adaptive platform trial would have assigned 34% more GBM patients to the effective arm of a trial relative to an RCT design.⁴ Alexander believed that adaptive platform trials had potential to usher in a new generation of effective treatments for GBM.

By mid-2017, Alexander and his colleagues had completed a master protocol for GBM AGILE; however, the research team faced several design and operational challenges as they prepared for the trial launch. For example, all trials needed a sponsor (i.e., the company or academic group leading the research effort); who would function as the sponsor of this trial? The *ad hoc* group that had come together to design the trial was not an operational entity, and no formal planning had yet occurred for how to build and sustain a structure to run the trial.

In addition, whereas traditional RCTs had clear beginning and end dates, adaptive platform trials were, in theory, indefinite in length. As new potential therapies emerged, adaptive platform trials could add treatment arms to study them in real-time using existing trial infrastructure. As an adaptive platform trial proceeded, arms with ineffective therapies were dropped, and promising therapies were “graduated” out of the trial for advancement through the formal regulatory approval processes. The vast majority of RCTs were funded by pharmaceutical companies, but Alexander knew that it was unrealistic to expect external funding from one sponsor to run an “evergreen” adaptive platform trial. How could he sustainably finance the launch and ongoing operations of GBM AGILE?

Glioblastoma

GBM was a rare, malignant, fast-growing cancer that originated in the supportive tissue of the brain.⁵ The National Cancer Institute estimated that in 2017, 23,800 cases of brain and central nervous system cancer would be diagnosed in the United States that year,⁶ of which about half would be GBM,⁷ and the cause of which was unknown.⁸ Men were diagnosed with GBM at a slightly higher rate than women, and most people were between the ages of 55 and 64 at the time of diagnosis.⁹ Clinically, GBM patients were categorized as either newly diagnosed (i.e., a first-time GBM diagnosis) or recurrent (i.e., a tumor had reappeared following an initial diagnosis and treatment). The standard of care and prognosis were different for each of these sub-types.

After diagnosis, physicians performed surgery to partially remove the tumor when possible. The tumor’s precarious location in the brain and its extensive, infiltrative nature usually prevented physicians from removing it entirely, but partial removal could alleviate some pressure on the patient’s brain. Physicians balanced the need to remove the tumor with the risk of affecting critical brain functions, such as speech and movement.¹⁰ After surgery, patients typically endured radiation and chemotherapy.¹¹ Despite these treatments, GBM recurred frequently, and prognosis was poor. Five-year survival rates (i.e., the likelihood that a patient would survive five years past diagnosis) were in the single-digits for some sub-types. As of 2013, only 10% of GBM patients were participating in clinical trials.¹²

Randomized Controlled Trials

In the 1940s, the scientific community began to rely on RCTs to assess the efficacy, safety, and appropriate dosage of drugs for diseases.¹³ RCTs, with their requisite two study arms, remained largely unchanged for the next 70 years.¹⁴

Researchers conducted RCTs to understand drug treatment effects for a wide range of conditions and diseases. For example, to test the efficacy of a new drug to lower LDL cholesterol levels against a standard treatment, an RCT enrolled a large number of eligible participants with high cholesterol. Researchers used a computer program to randomly assign these people to the treatment or the control group. Randomization, when properly conducted, resulted in a scenario in which patients' baseline characteristics such as age, gender, and starting LDL levels, along with unknown or unmeasured characteristics that may impact the outcome, were spread evenly, on average, across the two groups. At the end of the trial, researchers measured participants' LDL levels and assessed whether the treatment arm (patients receiving the new drug) had meaningfully different outcomes than the control arm (typically patients receiving the current standard treatment).¹⁵ (**Appendix A** shows some simple power calculations of the type used to determine the patient sample size required in a traditional RCT).

Before conducting an RCT, researchers developed a study protocol which specified the research question, the study design and methodology, inclusion criteria (i.e., a list of factors that determined whether someone was eligible to participate in the trial), and the study's primary endpoint(s), or the key outcome(s) of interest.¹⁶ It typically took several years to design and conduct an RCT.¹⁷

Frequentist vs. Bayesian Statistics

Traditional RCTs, which were based on the principles of "frequentist statistics," established study parameters at the beginning of the trial and held them constant throughout (e.g., fixed randomization to the two groups; analyzing data only once the trial ended). Because interpretation of the trial results was inextricably linked to the trial design, there was less flexibility in leveraging data that came in during the course of the trial.

By contrast, trials based on "Bayesian statistics" estimated the probability of treatment effects based on data as it accumulated, and researchers could use pre-specified algorithms to update probability estimates and alter the conduct of the trial in real-time.¹⁸ For instance, in a Bayesian trial, if accumulating data indicated that one drug outperformed another, researchers could use a pre-specified algorithm to incorporate this information into patient assignment. They could then move a higher proportion of new trial enrollees to the more promising therapy, a technique called "adaptive randomization." This type of dynamic adjustment of trial design would be considered taboo in frequentist trials.

As one scholar explained, "In the Bayesian approach, experiments can be altered in midcourse, disparate sources of information can be combined, and expert opinion can play a role in inferences."¹⁹ Trials rooted in Bayesian statistics often allowed researchers to more quickly identify potentially effective therapies.

Drug Approval Process in the United States

In the United States, pharmaceutical companies, biotechnology companies, medical technology companies, and academic research groups sponsored clinical trials to investigate potential new drugs. From 2006 to 2014, the proportion of trials financed by commercial companies increased by 43%, while the proportion funded by the U.S. National Institutes of Health (NIH), a government agency, declined by 24%.²⁰ As of 2013, the majority (85%) of clinical trials were funded by the pharmaceutical industry.²¹ Just over one million people participated in a clinical trial in the United States in 2013, of which 215,000 were cancer patients.²² Trial sponsors often struggled to recruit enough participants. One study found that up to 40% of cancer trials failed to enroll the necessary sample size.²³

To initiate a drug trial, the sponsor was required to file an Investigational New Drug (IND) application with the U.S. Food and Drug Administration (FDA), the agency responsible for regulatory approval of all medical products in the United States. Once this application was approved, drug trials typically occurred in three sequential phases (see **Exhibit 1** for detail on these phases).²⁴ Phase I trials were small, took place over the course of several months, and explored the safety and the appropriate dosage of new therapies. Phase II trials enrolled several hundred people, aimed to determine drug efficacy and side effects, and typically took up to two years. Lastly, phase III trials sought to confirm a therapy's efficacy, safety, and adverse reactions.²⁵ Significant time lags were common between phases of clinical research and some time was also needed for FDA review before and after each phase. All told, the process of moving a drug from the laboratory to the end user took approximately 12 years.²⁶

Trial sponsors typically outsourced the coordination of operational activities to contract research organizations (CROs), third-party companies that supported infrastructure of clinical trials. CROs collected fees from trial sponsors and then paid per-patient fees to one or more clinical trial sites where patients were given experimental drugs. Per-patient costs (used to manage the trial, pay employees, and conduct diagnostic testing) associated with new cancer drug trials in the United States were estimated to be \$59,500 or higher.²⁷

Once the trial sponsor believed it had garnered sufficient clinical evidence to demonstrate a new product's safety and efficacy, it filed a New Drug Application (NDA) with the FDA. Fulfilling its regulatory role, the FDA then reviewed the trial sponsor's evidence to decide whether to approve the new drug for use in humans.²⁸ Throughout its history, the FDA had approved nearly 1,500 drugs for prescription use.²⁹

Estimates of the cost of drug development ranged from \$868 million³⁰ to \$2.6 billion.³¹ Failure rates were high; just 10.4% of investigational drugs were eventually approved for use by the FDA.³² In addition, drugs found to be promising in phase II trials often performed poorly once they advanced to phase III; this was particularly true for cancer drugs.³³ One review of all phase III chemotherapy drug trials from 1998 to 2003 found that 81% of these trials showed lower response rates than their phase II counterparts.³⁴ Experts theorized that this was due to changing survival trends, newer imaging techniques, and differing clinical endpoints between trial phases.^{35,36} For GBM, particular problems included lack of randomization in phase II trials.³⁷

The Growing Importance of Biomarkers

The FDA was also paying increasing attention to evidence showing that specific sub-populations with certain observable characteristics, or "biomarkers,"^a were particularly responsive to certain drugs. In May 2017, Keytruda (pembrolizumab), Merck's FDA-designated "breakthrough" cancer immunotherapy drug, received FDA approval for use in all cancer patients with a specific biomarker.³⁸

^a A biomarker was a naturally occurring molecule or biological characteristic found in certain people that could influence drug treatment responses. As an example, a July 2017 study found that breast cancer patients with two biomarkers (cytoplasmic cyclin E and retinoblastoma protein) responded differently to first-line treatment than patients without these biomarkers. Source: University of Texas M.D. Anderson Cancer Center, "Study Shows Biomarkers Can Predict which ER-Positive Breast Cancer Patients Respond Best to First-Line Therapy," <https://www.mdanderson.org/newsroom/2017/06/study-shows-biomarkers-can-predict-which-er-positive-breast-cancer-patients-respond-best-first-line-therapy.html>, accessed September 2017.

In December of 2016, the FDA-NIH Biomarker Working Group published a detailed document, "BEST" (Biomarkers EndpointS, and other Tools Resource) to address the need to standardize language around the use of biomarkers in translational science and medical product development. Source: FDA-NIH Biomarker Working Group, "BEST (Biomarkers EndpointS, and other Tools Resource)," <https://www.ncbi.nlm.nih.gov/books/NBK326791>, accessed August 2017.

This was the first time the FDA had approved a drug to respond to a biomarker signature, rather than a specific cancer type (e.g., lung, colon, etc.).³⁹

Recent Developments in Clinical Trials

In the early 2000s, academics, medical practitioners, and statisticians began to question the dominant use of traditional RCTs and frequentist statistics in clinical research. While traditional RCTs maintained a high degree of rigor, critics contended that their rigidity deterred innovation and stalled improvements in treatment.⁴⁰ A small but growing community advocated for trials based on Bayesian statistics, which could expedite identification of effective therapies and reduce the cost of drug development without compromising the quality of trial results.⁴¹ The proliferation of computers with high-speed computational power (necessary for Bayesian analyses) enabled interested researchers to explore and experiment with Bayesian statistical methods.⁴²

Proponents of the Bayesian approach also pointed out that there was a mismatch between the evolving state of medicine and the inflexible parameters of traditional RCTs.⁴³ For example, in recent years, the scientific community had discovered that cancers were not nearly as similar as had historically been thought. Two patients who presented with lung cancer could have tumors with different characteristics and mutations. As the medical field advanced in its understanding of patient sub-populations with variants of the same disease, the number of possible drug uses multiplied. This was perhaps best appreciated in the context of personalized and precision medicine, where therapies could be targeted to the sub-population of patients most likely to benefit.^{44,b}

Yet the methodology for evaluating new therapies—traditional RCTs—remained the same. One scholar argued that traditional RCTs had become “straightjackets for clinical practice by providing answers to outdated questions.”⁴⁵ Further, the growth of personalized and precision medicine had led to a vast increase in the number of clinical questions to be answered, making the traditional RCT model—where each trial asked and answered only one question—seem inadequate.⁴⁶

Growing Use and Acceptance of Bayesian Trial Design

Throughout the 2000s, clinical trials and analyses based on Bayesian statistics had gained popularity. For example, of the 964 trials conducted at MD Anderson Cancer Center from 2000 to 2005, approximately 20% (195) included Bayesian analysis or design elements.⁴⁷ Regulatory bodies were also warming to the Bayesian approach. In June 2003, the cholesterol drug Pravigard Pac became the first drug approved by the FDA on the basis of a Bayesian efficacy analysis.⁴⁸ From 2000 to 2010, the share of medical devices approved by the FDA based on evidence from Bayesian analyses increased from none to about 5% to 10%.⁴⁹

The publication of guidance documents and other regulatory materials soon followed. In 2007, the European Medicines Agency (EMA), the body responsible for regulating drugs and medical devices in the European Union, published a reflection paper exploring methodological considerations for trials with an adaptive, Bayesian design.⁵⁰ In 2010, the FDA released draft guidance for using adaptive trials (based on Bayesian statistics) to evaluate drugs and biologics⁵¹ and in 2016, the FDA released official guidance for applying adaptive trial designs to medical device trials.⁵² These documents emphasized

^b Precision medicine was defined as “an emerging approach for disease treatment and prevention that takes into account individual variability in gene, environment, and lifestyle for each person.” Source: U.S. National Library of Medicine, “What Is Precision Medicine?” June 27, 2017, <https://ghr.nlm.nih.gov/primer/precisionmedicine/definition>, accessed June 2017.

that adaptive trials should establish strict control of type I error^c and plan for any potential trial modifications before, rather than during, the trial launch.

Scott Gottlieb, who became FDA commissioner in May 2017, was an enthusiastic supporter of the potential efficiency improvements in drug approval enabled by Bayesian statistics.⁵³ In his April 2017 confirmation hearing, Gottlieb said: “I think anything we can do to try to make [the drug approval] process more predictable, to create bright lines, to use better tools to evaluate safety and effectiveness that could bring down the cost while not doing anything to sacrifice . . . our ability to ferret out the safety of a product are things we should be looking at.”⁵⁴

Platform Trials, Adaptive Trials, and Adaptive Platform Trials

Platform Trials and Adaptive Trials

Both platform trials and adaptive trials were emerging clinical trial designs. They could be combined (as in GBM AGILE), but did not necessarily need to be used together.

Platform trials As compared with traditional RCTs – which tested one therapy in one treatment arm – platform trials simultaneously tested several therapies in multiple treatment arms.⁵⁵ Most often, these arms were compared to one common control arm. (See **Exhibit 2** for additional differences between a platform trial and a traditional RCT.)

Adaptive trials Adaptive trials were those that used a range of planned modifications to increase trial efficiency. While the definition of an adaptive trial differed among scholars and regulatory bodies (see **Exhibit 3**), some of the most commonly used adaptive elements included adaptive randomization, arm dropping, and seamless phase II/III transition, further explained below.

Adaptive Platform Trials

Adaptive platform trials combined multiple treatment arms – the hallmark of a platform trial – with elements of adaptive trial design. They also could incorporate other design features to highlight differences in treatment effects among individuals. For instance, these trials often assigned patients to treatment arms based on the presence of a specific characteristic or biomarker. This allowed researchers to evaluate, for example, whether a particular cancer therapy was more effective in individuals with a specific tumor expression.⁵⁶ Proponents of these trials argued that the ability to understand how therapies affected sub-groups was invaluable, especially in the context of precision medicine.⁵⁷ Adaptive platform trials could take on many different design features that led to efficiencies in the clinical research process. These included:

Adaptive randomization Adaptive randomization was a process that used information about the effectiveness of the therapies already under investigation to inform patient randomization.⁵⁸ For example, if one drug was found to be superior to the others among early enrollees in a trial, researchers could increase the proportion of new enrollees randomized into the more effective treatment arm. Participants would still be randomized, but the allocations would change to favor the superior treatment. As a result, a higher share of patients benefitted because they were exposed to a treatment

^c In statistical terms, a type 1 error occurred when a clinical trial made the mistake of concluding that the new therapy was superior to the placebo or standard of care when in reality, it was not. This was also sometimes called a “false positive.”

likely to result in better outcomes. Furthermore, researchers could more quickly confirm the estimated treatment effect because of the increased sample size in that particular study arm.

Arm dropping and “graduation” Another common adaptive trial feature was the ability to continuously add, drop, or “graduate” (progress) treatment arms throughout the life of the trial.⁵⁹ If new drug candidates emerged, researchers could add a treatment arm to an existing platform in order to study the effects of that particular drug. This had the appealing feature of allowing new candidate therapies to begin clinical trials quickly and efficiently by using existing research infrastructure. Concurrently, if a drug under investigation was found to be ineffective (either overall or in a specific sub-group of patients), researchers could drop that treatment arm (participants in that arm could then potentially be reassigned to other arms). Finally, if a drug in a given treatment arm was found to be effective, researchers “graduated” that therapy into the next phase of clinical trials.^{d,e}

Seamless phase II/III transition Seamless phase II/III transitions aimed to accomplish the goals of both a phase II clinical trial (i.e., further evaluate a drug candidate’s safety and efficacy) and a phase III trial (i.e., confirm efficacy) within the framework of a single adaptive trial.⁶⁰ These were designed to allow promising therapies to move straight from a registered phase II trial into a phase III trial following an interim analysis or based on accumulating data, thereby reducing the lengthy delays that often elapsed between the phases.⁶¹

Other novel trial designs for precision medicine Researchers had also developed other trial designs for studying targeted therapies and precision medicines, such as *umbrella trials* and *basket trials*. These trials were unique because of how researchers selected participants. Put simply, umbrella trials enrolled patients based on the presence of a shared disease type (e.g., colon cancer), regardless of individual patients’ particular mutations. Basket trials, by contrast, enrolled patients based on a shared mutation, regardless of their disease type. For example, two patients with the same mutation in a particular gene whose cancers presented differently (i.e., one had lung cancer and the other had colon cancer) would be grouped together in the same study (see Exhibit 4).⁶²

Standard RCTs were designed to answer one question. With biomarkers, however, each therapy now had multiple potential questions (patient biomarker types) to address. As Alexander explained, “When you think about a disease rather than a drug, there are obviously many questions – related to drugs, biomarkers, combinations, etc. – to be answered. Part of the move to platform trials is to ask and answer *many* questions about a disease or patient population and the Bayesian design supports that.”

Efficiency Gains from Adaptive Platform Trials

Each time a new trial began, regardless of its design, investigators invested time and money into developing a study protocol, hiring a CRO, recruiting a team of clinicians and statisticians, securing funding, obtaining ethical approval, identifying clinical trial sites, working with regulatory agencies, instituting standard operating procedures, establishing contracts, and enrolling participants. Because adaptive platform trials were ongoing, many of these “fixed costs” could potentially be shared across multiple treatment arms in a way that was not possible in traditional RCTs. Some investments, such as

^d A notable example was the I-SPY 2 trial of neoadjuvant experimental therapies in breast cancer, which identified two breast cancer therapies with biomarkers that had a high probability of success in phase III studies. Source: Alexander M. Castellino, “Two Drugs ‘Graduate’ From Novel I-SPY 2 Breast Cancer Trial,” *Medscape*, July 6, 2016, <http://www.medscape.com/viewarticle/865785>, accessed June 2017.

^e A design element closely related to graduation was **adaptive sample size**. In a traditional RCT, sample size estimates were made based on expected effect size (see **Appendix A**), but in adaptive trials, the sample size could be adjusted if the effect size turned out to be larger (or smaller) than anticipated.

CRO selection, ethical approval, and statistical design needed to occur only once, at the time the platform was launched. Thereafter, new drugs could be added to the trial continuously with lower start-up costs relative to what would be required to launch a new traditional RCT.

Adaptive platform trials also made it easier to test drugs for new indications by lowering the barriers to trial initiation and entry. For example, if an existing drug was known to be effective for one cancer with potential benefit in other cancers, its manufacturer could add the drug to an existing adaptive platform trial for a different indication (or a different patient sub-group than the existing indication(s)). As a result, more treatments might be identified for indications that otherwise would not have been pursued. Increasing the number of disease-drug pairs being studied in clinical trials would address an ongoing challenge in cancer research. As articulated by one scholar: “We speak of false negatives and false positives, but both are dwarfed by false neutrals – therapies that have not been and may never be evaluated in clinical trials.”⁶³

The most well-known trial of this type to date, I-SPY 2, compared the effects of treating breast cancer with chemotherapy alone (the standard of care) versus chemotherapy plus biomarker-specific drugs prior to surgery.⁶⁴ Launched in 2011, I-SPY 2 evaluated whether the treatment eradicated the patient’s tumor, known as a pathologic complete response (the trial’s endpoint).⁶⁵ As of June 2017, I-SPY 2 had tested drugs from Abbott, Abbvie, Amgen, Genentech, Medivation, Merck, Plexxicon, and Pfizer/Puma.⁶⁶ The trial had evaluated 12 therapies, five of which had advanced for additional evaluation.⁶⁷

I-SPY2 had also shown that adaptive platform trials could save a significant amount of time in set-up. It took I-SPY 2 researchers roughly five months to move from negotiations with drug companies to patient enrollment into the trial, whereas an RCT typically took 18 to 36 months to do so.⁶⁸ I-SPY2 relied on multiple sources of funding including research grants from the FDA, NIH, academic research centers, and patient advocates, in addition to support from biopharmaceutical manufacturers. Researchers had also designed adaptive platform trials to study a range of other conditions and diseases, including antibiotic resistance, dementia, influenza, pancreatic cancer, and pneumonia.^f (See **Exhibit 5** for a list of related trials and their associated characteristics and design features.)

An Adaptive Platform Trial for GBM?

Alexander, who was an Associate Professor at Harvard Medical School in addition to his work at Dana-Farber, first became enthusiastic about the efficiency gains associated with adaptive platform trials in the early 2010s. “If you look at the landscape of GBM development over the last 10 to 15 years,” Alexander said, “it’s amazing how many single-arm, phase II trials appeared positive but were found to be negative in phase III. If we can design an adaptive platform trial that can seamlessly transition effective GBM therapies from a more rigorous phase II to phase III, that would go a long way toward identifying or disproving new therapies more quickly.”

GBM, Alexander thought, was particularly well-suited to be studied in this setting. As he explained, “An adaptive platform trial gains efficiencies through the ability to quickly gather and use meaningful information. With GBM, because the prognosis is so poor, we can collect and use information about

^f Notably, in 2014, researchers designed an adaptive platform trial to study treatments for the Ebola virus in West Africa. Because the epidemic slowed in early 2015, the trial was never launched; however, this demonstrated that the applications of adaptive platform trials could extend to both communicable diseases and a developing country context. Source: Scott M. Berry et al., “A Response Adaptive Randomization Platform Trial for Efficient Evaluation of Ebola Virus Treatments: A Model for Pandemic Response,” *Clinical Trials* 2016, Vol. 13(1):22-30.

how treatments affect patient survival in real-time to inform the trial.” In addition, given the relatively low prevalence of GBM, pharmaceutical companies had weaker economic incentives to develop drugs specifically targeted at that cancer. Indeed, research had shown repeatedly that smaller drug markets attracted fewer pharmaceutical entrants.^{69,70} As Alexander noted, “A platform trial provides a turn-key way for companies to look at another indication for an existing drug. For diseases that don’t receive a lot of individual attention from pharmaceutical companies like GBM, this kind of trial can really open the door for testing a range of new therapies.”

In 2012, Alexander co-authored an article with statistician Lorenzo Trippa on the potential to use adaptive trial designs for GBM. The following year, he gave a presentation on the same topic at the American Society of Clinical Oncology conference. “I was really nervous about that presentation,” he recalled, “because Don Berry, who is basically the Godfather of Bayesian trial designs, was the discussant.” Alexander was pleased to discover that Berry was approachable and eager to discuss his ideas around more efficient trial design for GBM. They talked at length after the presentation and stayed in contact about the topic. In 2013, Alexander and Berry jointly published an article exploring how adaptive platform trial methodology could be applied to GBM.⁷¹

Additionally, with funding from a Burroughs Wellcome Innovations in Regulatory Science Award, Alexander and his colleagues at Dana-Farber, including Trippa, neuro-oncologist Patrick Wen, and pathologist Keith Ligon, had designed the INDividualized Screening Trial of Innovative Glioblastoma Therapy trial (INSIGHt; NCT02977780).⁷² INSIGHt was designed to be an efficient phase II screening adaptive platform trial to identify promising therapies in association with genomic biomarker analysis for patients with a particular indication—namely those with newly diagnosed GBM that tested as “unmethylated” for the MGMT biomarker. INSIGHt eventually launched in January 2017 with the support of the Accelerate Brain Cancer Cure (ABC²) foundation along with the pharmaceutical companies Puma, Eli Lilly, and Celgene, whose drugs were being tested in the trial.

In parallel, Berry had invited Alexander in 2014 to join a group called GBM AGILE that was expressly created to imagine, design, and write a protocol for a novel global clinical trial for GBM. The group was led by Dr. Anna Barker, the director of transformative healthcare knowledge networks at Arizona State University’s National Biomarker Development Alliance, and the trial would be led by Principal Investigator Timothy Cloughesy, the head of neuro-oncology at the University of California, Los Angeles. Barker was an expert in understanding the needs of targeted cancer research, having been the Deputy Director of the National Cancer Institute (NCI) and Deputy Director for Strategic Scientific Initiatives, at the time that it launched the Cancer Genome Atlas Project.⁷³ The GBM AGILE group also included a collaborative group of oncologists, pathologists, radiologists, and statisticians hoping to design an adaptive platform trial for GBM.

Over the next year, this group met routinely to discuss the particulars of the trial, including biomarkers to explore, promising drugs for GBM, and possible endpoint selections. As Barker remarked, “The concept of combining our knowledge and expertise to create the first global adaptive trial for GBM quickly became something of a ‘movement.’ . . . United by a simple goal of bringing better treatments to GBM patients, everyone stepped out of their silos, left their egos, and paid their own way [to work on GBM AGILE].”⁷⁴ Meredith Buxton, director of clinical trial strategy at Berry Consultants, a well-known consulting firm founded by Don Berry that specialized in statistical support for adaptive trials, reflected on the uniquely collaborative nature of the broader community around adaptive platform trials:

These trials are often spurred on by a small but committed group of clinicians, researchers, and advocates who want and expect better for patients and for the

advancement of research. However, for these efforts to really take hold, a larger consortium of academic investigators, industry partners, advocates, funders, and regulators is needed. Success lies in moving from an idea generating phase at a 50,000-foot level to the on-the-ground implementation phase; it is crucial to strike a balance between what is scientifically valuable and operationally feasible. This is best accomplished through strategic partner alliances and a consortium of committed and active collaborators to guide important components of the trial.

Designing the GBM AGILE Trial

By 2015, following discussions with the FDA, the GBM AGILE group had developed a master protocol (i.e., a document that contained all of the governing rules for conduct and analysis) for the phase II trial. Alexander and his colleagues wrote that the protocol represented “the efforts of over 130 oncologists, statisticians, pathologists, neurosurgeons, imagers, and translational and basic scientists” from around the world.⁷⁵ The research group planned to use overall survival as the trial’s endpoint. Because GBM survival rates were poor, the trial would relatively quickly detect any survival benefits from new drugs.

To strike a balance between trial efficiency and scientific rigor, Alexander explained, “Our trial design team, led by Berry, worked with the FDA to figure out how we can best use the science at our disposal to avoid type 1 error but also get superior treatments to patients.” GBM AGILE was originally designed to use Bayesian adaptive randomization, increasing the probability that patients would be assigned to a trial arm (therapy) from which they were more likely to benefit.⁸ But Alexander was reminded of hesitations he had felt after an earlier conversation with Richard Pazdur, the director of the FDA’s Oncology Center of Excellence, about the challenges of using different endpoints for various approval pathways. “What if we showed an increase in overall survival—a meaningful trial endpoint—but with this new kind of design?” Alexander wondered. “Would we be stuck doing another separate phase III trial to confirm the results? That didn’t seem to be in the best interest of patients.”

The GBM AGILE team and the FDA ultimately resolved this issue by employing an innovation from Berry based on a two-stage randomization design. The first stage would consist of the already-planned adaptive randomization. Therapies that showed compelling improvements in survival in this stage would then seamlessly transition to a second stage of tests using standard, fixed randomization to confirm the findings. This would support trial registration, an important component for regulators like the FDA and EMA. Stage 1 would “learn” – both about the efficacy of new therapies and the right biomarker-defined population to use them. Stage 2 would then “confirm.” In characteristic platform trial fashion, new therapeutic arms could be added to the trial over time while other arms completed testing (see **Exhibit 6** for additional detail on trial structure).

The design of GBM AGILE was a significant step forward. By including newly diagnosed and recurrent patients in the same trial, there was potential to make clinical research and clinical practice more seamless while providing efficiencies. As Alexander explained, the two-stage design created to support ultimate drug approval was a complete solution to optimal development from the beginning of the traditional phase II through the end of phase III.

⁸ The earliest patients in the first stage of the trial would be randomly assigned (in pre-specified proportions) to treatment arms. Once data began to accumulate, information from interim analyses could be used to *adaptively* assign subsequent incoming patients to treatment arms.

Operationalizing the Trial

To turn the design and protocol into reality, Alexander and his colleagues needed to identify a sponsor with an appropriate governance and operational structure to oversee the trial. An evergreen adaptive platform trial for GBM created an opportunity to establish an entity that was explicitly focused on developing therapies *for* patients with GBM, a mission and focus that no other player in the therapeutic development space had. “This was a really exciting feature of an ongoing organization to me,” Alexander said, “given the potential to change the therapeutic and biomarker development process to be more patient-centric.”

There were practical considerations as well. “At the beginning, we had no mechanism to engage with pharmaceutical companies,” Alexander recalled. “We started to talk to them, but when we reached the point of negotiating budgets and contracts, it became clear that we needed to transition from the design phase to the operational phase.” Thus in April 2017, Alexander and his colleagues decided to found a non-profit organization that would sponsor the GBM AGILE trial called the Global Coalition for Adaptive Research (GCAR). Alexander would ultimately agree to serve as CEO and president of the GCAR as the group moved to identifying, vetting, and securing contractual relationships with vendors, legal counsel, and identifying a CRO to partner with in creating and running GBM AGILE.

By the summer of 2017, the GCAR had also finalized a contract with Berry Consultants (see **Exhibit 7** for a list of the services that this group specialized in providing). Berry Consultants would coordinate the statistical strategy of the trial. The trial would take place across multiple clinical sites (e.g., university research centers, hospitals, and cancer institutes) throughout Australia, China, and the United States. GBM AGILE would enroll both newly diagnosed and recurrent GBM patients and evaluate which therapies worked best for each type. Alexander remarked:

This is a really unique approach. Usually, pharmaceutical companies test therapies on recurrent GBM patients first—almost like the equivalent of metastatic (or later-stage) disease for other types of cancers. If the drug shows promise among this population, the company might then test it on newly diagnosed patients. GBM AGILE enrolls both newly diagnosed and recurrent patients. This gives us more statistical power because our sample size is larger. The approach also starts to dissolve the machinery of clinical trials. You could imagine that a patient comes into an oncology clinic and is enrolled on GBM AGILE simply as part of his or her GBM care because all patients are eligible. As a result, we’re generating information from as many patients as we can. We would truly be transforming clinical care into an optimized learning environment where patients had access to the latest, more personalized therapies.

Alexander knew that an early priority would be to create a sustainable financing model for the GBM AGILE trial. It was clear that the GCAR could not request external grant funding for a GBM trial of indefinite size and duration from any of the typical clinical research funders in the pharmaceutical industry. The GCAR would coordinate the trial, but there would be substantial upfront financial and coordination costs. Alexander and his colleagues needed to resolve several potential financing questions—both in the short-run as well as in the longer run, once the trial was underway.

In the short term, the GCAR needed a plan for launching the trial without an upfront source of money. Who would provide the necessary funds and when? To what extent could GBM AGILE expect to later receive government grants and academic support? What role would patient advocacy and philanthropic organizations play in supporting the trial? How would contracts with pharmaceutical

companies look? And might these contracts change as the trial evolved and became established in the clinical research world? They considered their options.

A clear option would be to have pharmaceutical companies fund individual arms of the trial. If a manufacturer wanted to include one of its drugs in the GBM AGILE trial, it would pay the GCAR to run a trial arm. In such a case, it was envisioned (but not necessary) that the manufacturer would retain the rights to any successful NDA eventually granted by the FDA. The GCAR would need to charge these companies a high enough amount to both cover overhead costs as well as per-patient fees at its trial sites and support the GCAR's operations. "Trials run by cooperative groups typically pay much less per patient than pharmaceutical companies," said Alexander. "And trial sites have to recover enough funds to support their own research infrastructure." The GCAR needed to pay sites an attractive enough fee per patient to allow GBM AGILE to be financially sustainable at the site level. Furthermore, in order for the trial to be a success, pharmaceutical partners would have to be willing to include their most promising new drugs in GBM AGILE.

Alexander and his colleagues faced another unique challenge as they initiated the trial. Once the trial was operational, costs would be shared across all the treatment arms. But at the trial's outset, no arms yet existed to bear costs. As Alexander said, "With a traditional RCT, you could create an all-encompassing budget for the National Cancer Institute or a pharmaceutical company. GBM AGILE is a totally different animal. We won't have a complete budget to draw from when we initiate the trial. As such, it requires a completely new partnership model."

To solve this challenge, a strategy was devised to ask trial partners to defer payment from the GCAR until trial arms began to be funded. Alexander elaborated, "The trial design presents an interesting opportunity for our partners to stake their claim as the 'go-to' resources for overseeing adaptive platform trials. If the broader research field moves in this direction, our partners will be well-positioned to pick up contracts. I want all of us to have a stake in the trial's success." Melissa Paoloni, director of clinical science strategy at Berry Consultants, reflected:

The goal is to create a model with a clear value proposition for all stakeholders. It would be a challenge for an individual pharma partner to replicate the innovative infrastructure and shared efficiencies, which create both equity and efficiency for all partners. This approach—which is applicable to the adaptive platform trial space in general—lowers the bar for bringing new products into a trial and therefore broadens the pipeline of drugs that can be explored. Coupled with a stronger scientific basis and valuable data, this design creates value for all involved—the research community, industry partners and, most importantly, the patients.

Other Financing Opportunities

The novel setup of GBM AGILE also created a number of other business opportunities. One option was to pursue a licensing or royalty-sharing arrangement with smaller biotechnology companies. Many of these companies, Alexander explained, did not have the resources to conduct their own phase II or phase III trials in GBM. Often, their only pathway to market was to license their drugs to a larger pharmaceutical company that would then fund clinical trials. Alexander explained, "We could imagine an arrangement where some of these companies license their drugs to GBM AGILE instead, and we run a trial arm to test the company's drug [in GBM]. Or, they could enter into a royalty-sharing arrangement with us where we test the drug through our trial platform. If it is found to be effective and eventually makes it to market, the company shares some of its revenues with the GCAR." The GCAR could then use cash flows from licensing or royalty-sharing to support its ongoing operations

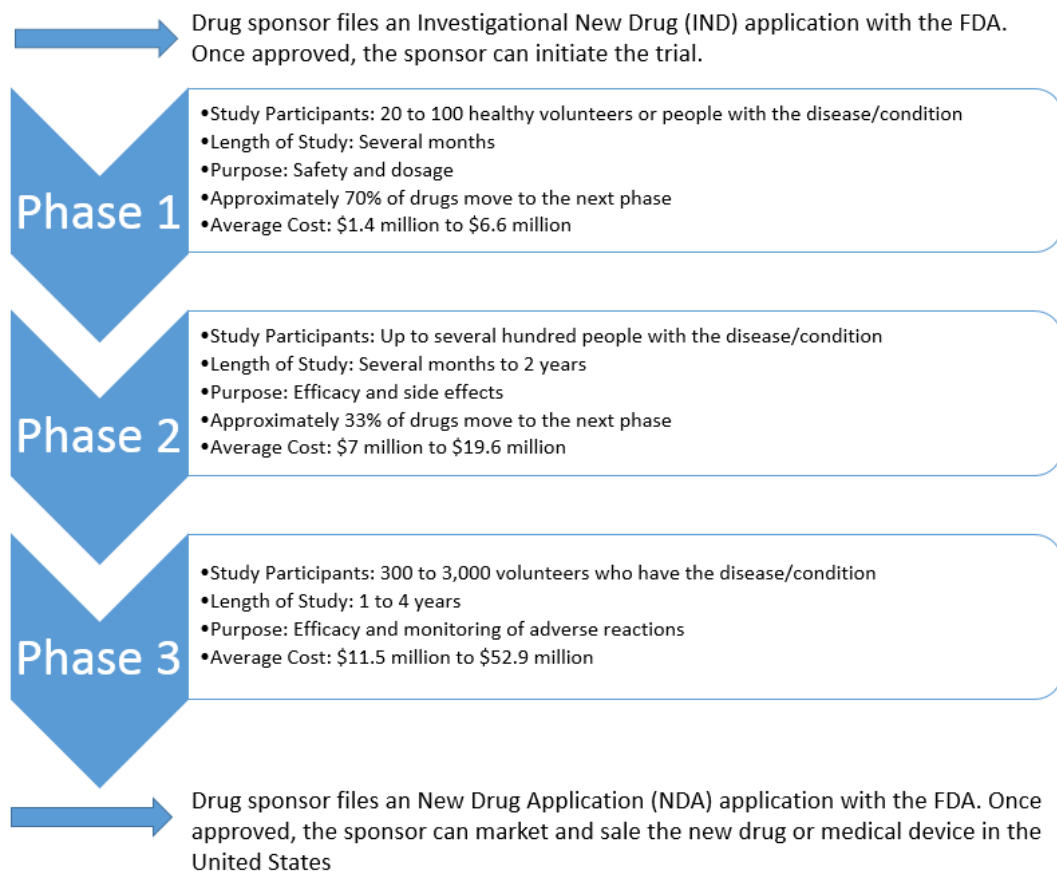
and/or to fund additional trial arms. Such an arrangement, Alexander explained, was not limited to smaller biotechnology companies. Major pharmaceutical companies might also have interest.

Another option was to monetize the data generated by GBM AGILE. “Both pharmaceutical companies and academic institutions will find the data valuable,” explained Alexander. He wanted academic researchers to have lower-cost access to the data than for-profit companies, given how important biomarker data would likely be in advancing precision medicine and cancer research. As he expressed, “We could possibly have two subscription models for accessing the data: one at a premium price for biopharmaceutical companies and the other free or at a discounted price for academic researchers.”

The GCAR could also consider developing biomarkers for predicting treatment efficacy or dosing recommendations for payers. Health insurance companies were increasingly interested in the potential to use biomarkers in coverage decisions. For example, knowing which patients were likely to have a very positive or very negative reaction to a therapy or which doses were most likely to be effective for a given patient would allow an insurer to make a value-driven decision about reimbursement for a given treatment and patient.⁷⁶ Payers were already discussing the potential for biomarkers to help drive coverage decisions.⁷⁷ Alexander expected other parties in the health care system to have an increased interest in developing biomarkers over the coming years, and in theory, the GCAR would be well positioned to be a service provider in this space.

A final idea was to have the GCAR act as a coordinator for all of a patient’s GBM care, somewhat like a managed care organization—a group of health providers that coordinate patient care—for GBM. Alexander pointed out, “We are already paying for all of these experimental therapies and specifically outlining a care management plan anyhow; why not pay for and manage the standard treatments patients will be receiving from the same oncologists?” None of the funding options were mutually exclusive, but the GCAR would need a plan for moving forward.

More broadly, decisions the GCAR made would help advise other organizations focused on finding better treatments for cancer and other diseases. For instance, the Multiple Myeloma Research Foundation (MMRF) would be closely monitoring the ways in which Alexander and the GCAR co-founders financed the GBM AGILE trial. As Kathy Giusti, founder of the MMRF and co-chair of the Kraft Precision Medicine Accelerator at Harvard Business School, remarked, “Innovating in the clinical trial design space is of crucial importance to improving outcomes for cancer patients. Everyone benefits when leaders work across cancers to design and support innovative trials conducted by trusted third parties.”

Exhibit 1 Phases of the U.S. FDA Screening Process for New Drugs

Source: Casewriter, based on information from: U.S. Food and Drug Administration, "Step 3: Clinical Research," May 25, 2017, <https://www.fda.gov/forpatients/approvals/drugs/ucm405622.htm>, accessed June 2017.; Cost Data from: Aylin Sertkaya, Hui-Hsing Wong, Amber Jessup, and Trinidad Beleche, "Key Cost Drivers of Pharmaceutical Clinical Trials in the United States," *Clinical Trials* 2016, 13(2): 117-26.

Exhibit 2 Differences between a Traditional RCT and a Platform Trial

Characteristic	Traditional Trial	Platform Trial
Scope	Efficacy of a single agent in a homogeneous population	Evaluating efficacy of multiple agents in a heterogeneous population; explicitly assumes treatment effects may be heterogeneous
Duration	Finite, based on time required to answer the single primary question	Potentially long-term, as long as there are suitable treatments requiring evaluation
No. of treatment groups	Prespecified and generally limited	Multiple treatment groups; the number of treatment groups and the specific treatments may change over time
Stopping rules	The entire trial may be stopped early for success or futility or harm, based on the apparent efficacy of the single experimental treatment	Individual treatment groups may be removed from the trial, based on demonstrated efficacy or futility or harm, but the trial continues, perhaps with the addition of new experimental treatment(s)
Allocation strategy	Fixed randomization	Response-adaptive randomization
Sponsor support	Supported by a single federal or industrial sponsor	The trial infrastructure may be supported by multiple federal or industrial sponsors or a combination

Source: Scott M. Berry, Jason T. Connor, and Roger J. Lewis, "The Platform Trial: An Efficient Strategy for Evaluating Multiple Treatments," *JAMA* 2015 313(16): 1619-20.

Exhibit 3 Varying Definitions of Adaptive Trials

Adaptive Trials

FDA: “An *adaptive design clinical study* is defined as a study that includes a prospectively planned opportunity for modification of one or more specified aspects of the study design and hypotheses based on analysis of data (usually interim data) from subjects in the study.”

European Medicines Agency: “Adaptive pathways is a scientific concept for medicine development and data generation which allows for early and progressive patient access to a medicine. [. . .] Adaptive pathways is based on three principles: iterative development, [. . .] gathering evidence through real-life use to supplement clinical trial data; [and] early involvement of patients and health-technology assessment bodies in discussions on a medicine's development.”

2005 Article: “We will refer to an adaptive design as a design that allows modifications to some aspects (e.g., trial procedures and/or statistical procedures) of an on-going clinical trial after its initiation, without undermining the validity and integrity of the trial.”

2006 Article: “By *adaptive design* we refer to a clinical study design that uses accumulating data to decide how to modify aspects of the study as it continues, without undermining the validity and integrity of the trial.”

2015 Article: “[Adaptive trials are those] in which unblinded data are monitored and used to determine the future course of the trial based on prospectively defined decision rules.”

2017 Article: “Adaptive Designs (ADs) use data accumulated at interim time-points in the study to allow to modify elements of the trial without increasing bias, or undermine the validity of the trial results or the integrity of the trial itself.”

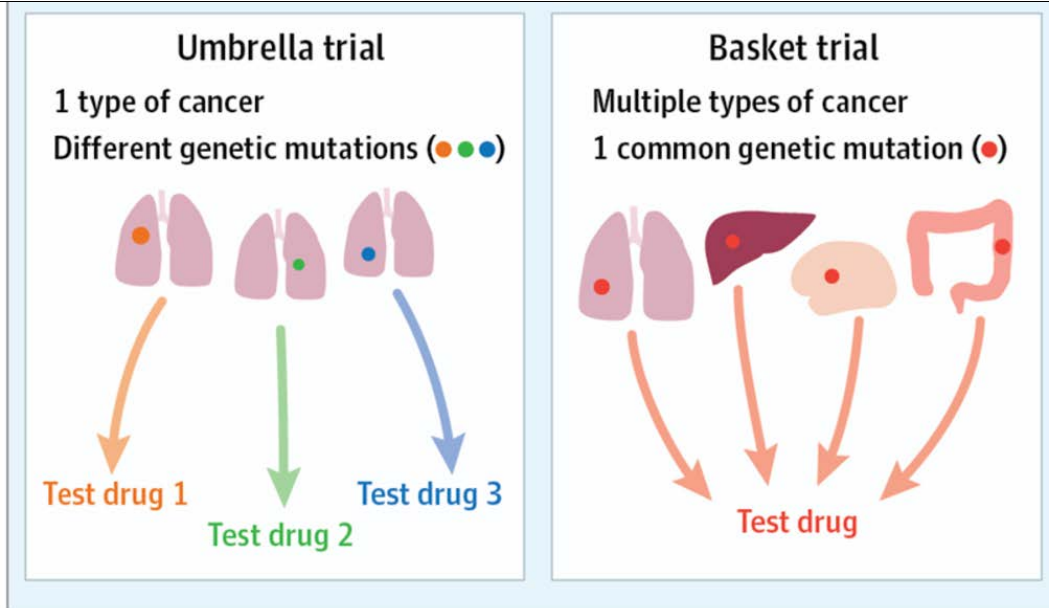
Source: U.S. Food and Drug Administration, “Adaptive Design Clinical Trials for Drugs and Biologics,” February 2010, <https://www.fda.gov/downloads/drugs/guidances/ucm201790.pdf>, accessed June 2017.; European Medicines Agency, “Adaptive Pathways,” http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000601.jsp, accessed June 2017.; Stein-Chung Chow, Mark Chang, and Annpey Pong, “Statistical Consideration of Adaptive Methods in Clinical Development,” *Journal of Biopharmaceutical Statistics*, (2005) 15(4): 575-91.; Paul Gallo et al., “Adaptive Designs in Clinical Drug Development – An Executive Summary of the PhRMA Working Group,” *Journal of Biopharmaceutical Statistics*, (2006) 16(3): 275-83.; Donald A. Berry, “The Brave New World of Clinical Cancer Research: Adaptive Biomarker-Driven Trials Integrating Clinical Practice with Clinical Research,” *Molecular Oncology* 9(2015): 951-9. Jianchang Lin and Veronica Bunn, “Comparison of Multi-Arm Multi-Stage Design and Adaptive Randomization in Platform Clinical Trial,” *Contemporary Clinical Trials* (2017) 54: 48-59.

Exhibit 4 Sub-Types of Platform Trials: Umbrella and Basket Trials

Sub-Types of Platform Trials

Umbrella trial: "enrolls patients who share the same basic cancer type, performs molecular marker testing for a wide array of potential targets, then assigns patients to an arm of the study based on the presence of a mutation matched to a potentially effective treatment for that marker."

Basket trial: "enrolls patients who have the same genetic mutation, whether their cancer originated in the lung, breast, colon, liver, or any other organ, then has patients all receive the same novel treatment that targets that specific marker."



Source: Jack West Howard, "Novel Precision Medicine Trial Designs: Umbrellas and Baskets," *JAMA Oncology* March 2017 3(3): 423.

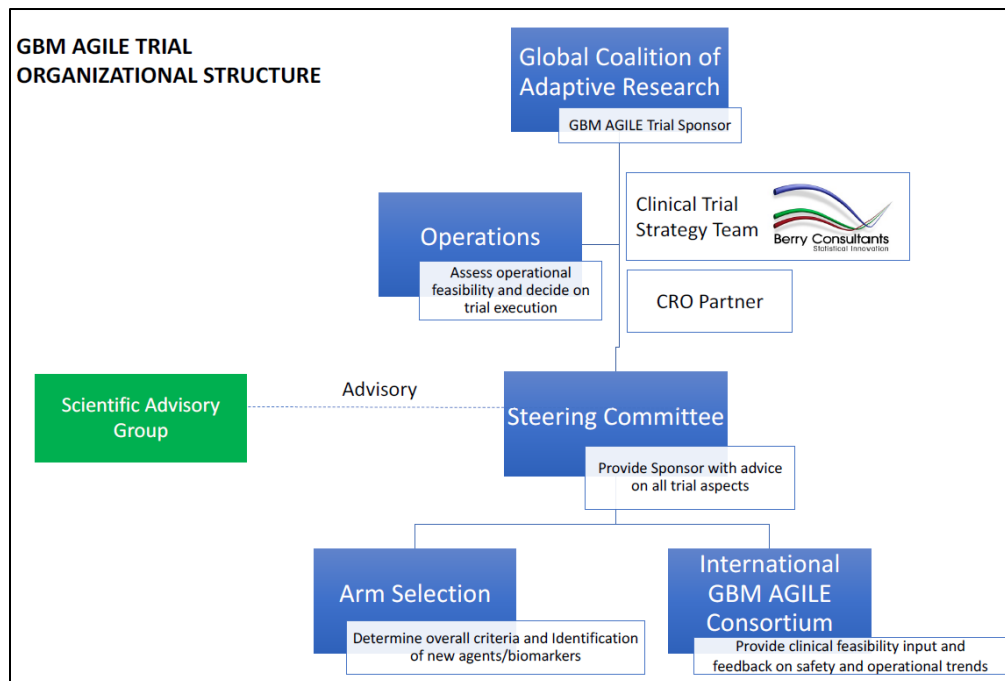
Exhibit 5 Adaptive Characteristics of Selected Trials

Trial	Patient Selection	Study Arms			Within-trial Adaptations			Integration
	Sub-type Stratification	Multiple Arms	Multiple Domains	Common Control	RAR	Interim Frequency	Staggered Arms	Bayesian Model
I-SPY 2 (Breast Cancer)	Yes	Yes	No	Yes	Yes	2 weeks	Yes	Yes
EPAD (Alzheimer's Disease)	Yes	Yes	No	Yes	No	Quarterly	Yes	Yes
DIAN (Alzheimer's Disease)	No	Yes	No	Yes	No	Biannually	Yes	Yes
Precision Promise (Pancreatic Cancer)	Yes	Yes	No	Yes	Yes	Monthly	Yes	Yes
PREPARE FLU (Influenza)	Yes/No	Yes	No	Yes	Yes	Weekly	Yes	Yes
REMAP-CAP (Pneumonia)	Yes	Yes	Yes	Yes	Yes	Monthly	Yes	Yes
INSIGH T (Glioblastoma)	Yes	Yes	No	Yes	Yes	Monthly	Yes	Yes
GBM AGILE (Glioblastoma)	Yes	Yes	No	Yes	Yes	Monthly	Yes	Yes

Source: Casewriter, adapted from: Derek C. Angus and Brian M. Alexander, "Adaptive Platform Trials: Definition, Design, Conduct and Reporting Considerations," presentation, *The Adaptive Platform Trials Coalition*, May 2017.

Note: RAR = Response-adaptive Randomization.

Exhibit 6 GBM AGILE Trial Organizational Structure



Source: Company documents.

Exhibit 7 Berry Consultants Background and Service Offerings

CLINICAL TRIAL STRATEGY TEAM OVERVIEW

To shorten the clinical development process by designing and implementing innovative, platform, and adaptive trials across all clinical areas.

GOALS AND OBJECTIVES:

- To provide strategic, scientific, and process design and operational expertise to groups wishing to establish large-scale clinical trial platforms for drug, device or diagnostics evaluation.
- To support the implementation and improvement of innovative (such as Bayesian), platform, oncology, other clinical trials.
- To foster the adoption of large-scale clinical trials by providing expert training to groups to execute such trials.



CTS SERVICES OFFERINGS

PRIMARY SERVICES:

- Design and Develop Innovative Trials
- Create Clinical Development Strategy
 - Consortium Building, Partner Alliances, and Regulatory Strategy
- Construct Trial Components and Processes
 - Recruitment Strategy Data Management Plan, Safety Management Plan, Drug Acquisition and Management
- Implement and Operationalize Large-Scale and Complex Trials
 - Oversight to include financial, contractual, IT, CRO, other vendors

ADDITIONAL SERVICES:

- Consulting Services
- Training /Education



“PAIN POINTS” TO FOCUS PLATFORM DEVELOPMENT

- Consortium Building
- Partnership Alliances
- Regulatory Strategy
- Recruitment
- Data Management
- Safety Management
- Drug Acquisition and Management
- Trial Oversight (financial, contractual, IT, CRO, other vendors)
- Training/Education



Source: Berry Consultants, used with permission.

Appendix A Simple Power Calculations

Traditional RCTs are designed to detect pre-specified outcome(s) of interest in the study population. In a traditional frequentist study design, researchers are interested in hypothesis testing, which has the primary goal of testing and potentially rejecting the null hypothesis, or demonstrating that the novel treatment has measurably different outcomes, on average, compared to the control arm of the study.

A critical element of study design is the power calculation. An accurate power calculation ensures that researchers enroll a sufficient study population to reject the null hypothesis if the novel treatment is, in fact, measurably different from the control treatment. Three inputs determine the required sample size: the expected effect size, the chosen significance level, and power (**Table A-1**).

Table A-1: Determinants of Sample Size

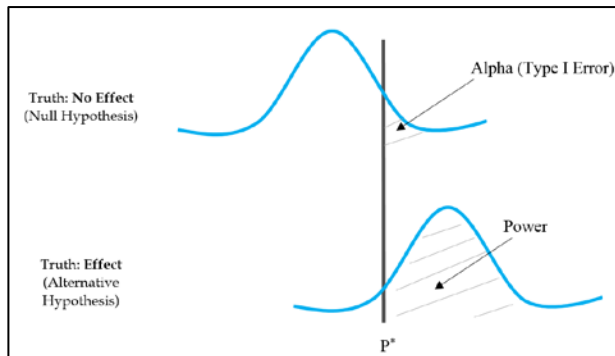
Input	Typical Values
Effect size of the alternative treatment	Context-specific
Significance level	0.05; 0.10
Power	0.8-0.95

The **effect size** is the expected difference in outcome(s) between the treatment and the control arm. For example, do researchers expect the new drug to perform marginally better than the existing treatment? Or do they anticipate double or triple the effect size? The answer to this question will affect sample size. Ideally, researchers estimate effect size by looking at previous research, such as a pilot study or observational study, and/or by taking into account the effect sizes considered clinically meaningful by experts or regulators, such as the U.S. Food and Drug Administration or the European Medicines Agency. Overestimating the treatment effect might lead to an *underpowered* study that incorrectly fails to reject the null. But underestimating treatment effect might lead to an *overpowered* study that enrolls more patients than necessary to answer the question.

The **significance level** and **power**, explained below, are both statistical parameters that describe the degree of certainty that accompanies the study's conclusion(s).

The **significance level** (also called alpha or type I error) is the probability of rejecting the null hypothesis when the null is true (i.e., the novel treatment is not different from the control, but researchers *incorrectly* conclude that it is). When studies refer to their results as "statistically significant at the 0.05 (or 5%) level," the study has a 5% type I error (false positive) rate, or an alpha of 0.05. If 100 studies are conducted that all have a 0.05 alpha level, in expectation, five studies will erroneously reject the null.

Power is the probability of *correctly* rejecting the null (i.e., identifying an effect when there is a true effect). The probability of type II error (a false negative) and the power of a test sum to 1. Therefore, the probability of a false positive is 1 minus the test's power. **Figure A-1** below shows the probability regions as they relate to the null hypothesis compared to a hypothetical alternative sample distribution.

Figure A-1 Probability Regions as Related to the Null and Alternative Hypotheses

With fixed values of the treatment effect, significance level, and power, researchers can calculate the required sample size of the study population. In a simple t-test (see **Equation A-1**), which tests the difference between two sample means, the effect size is calculated as the difference between the two sample groups' means divided by the common standard deviation. Notably, in cancer trials, power calculations are traditionally based on overall response rate (ORR), progression free survival (PFS), or overall survival (OS), rather than effect size.

Equation A-1 Simple t-test

$$effect\ size = \frac{|\mu_1 - \mu_2|}{\sigma}$$

Suppose that researchers are testing a new therapy for lowering elevated LDL cholesterol levels. Based on earlier drug trial phases, they have data that suggests the following:

Table A-2 Sample Values for a Cholesterol Therapy

	Estimate (mg/dL)
μ_1 : Placebo mean LDL	150
μ_2 : Novel therapy mean LDL	160
σ : Overall standard deviation in LDL across both groups	20

With these estimates, the target effect size would be 0.5 (i.e., $|\mu_1 - \mu_2|/\sigma$ or $|150-160|/20$). As stated previously, the target effect may be derived from earlier studies or may be the smallest effect that would demonstrate a meaningful clinical contribution given the context. **Table A-3** shows the required sample size based on common combinations of significance levels and power.

Table A-3 Required Sample Size based on Power and Significance Estimates

		Significance-level		
		0.1	0.05	0.01
Power	0.8	51	64	96
	0.9	70	86	121
	0.95	88	105	145

*Note these are sample sizes for each treatment group. The overall study is twice this size. It is also possible to create control and treatment arms which differ in size by specifying a ratio of treatment to control patients.

Calculating a sample size for a study is, in many cases, an automated process. Researchers can go online and find a sample size calculator (for example, there are resources available via statistical software packages such as STATA and R), and they can always simulate study characteristics and sample sizes for more complicated study designs. What is more important is understanding the influence that inputs have on sample size. For instance:

- As the effect size decreases, the space separating the two bell-curves in **Figure A-1** decreases, requiring a larger sample size to estimate the treatment effect.
- As the significance level decreases, the vertical line that crosses the two curves moves to the right, decreasing power (unless the sample size is simultaneously increased).
- As the variance in the outcome measure increases, the two curves in **Figure A-1** spread, increasing overlap. This makes it harder to detect an effect, requiring a larger sample size.

What Can Go Wrong

Although a sample size calculation itself is relatively straightforward, there can be underlying issues with the study that can undermine the validity of the calculation. For instance, the minimal detectable effect is often uncertain before the study has been conducted and researchers' estimates may be overly optimistic. If the true effect is smaller than the pre-specified effect, this may result in an "under-powered study."

Although this type of mistake may seem avoidable, the smallest meaningful effect is often hard to define in advance, and budgetary constraints or other real-world constraints or logistical issues may restrict the sample size in ways that lead to an under-powered study. Using ill-suited statistical methods can also be an issue. If the outcome is not distributed normally, for example, using a statistical test that assumes normality, such as the effect size calculation in **Equation A-1**, will generate inaccurate results.

Note: One of the statistical advantages of an RCT is the simplicity of methods like two-sample t-tests. But in many clinical cases, researchers are interested in different outcomes, like time-to-event analyses (i.e., the length of time to a given outcome, such as all-cause mortality or hospitalization). Time-to-event analyses require a different type of sample size calculation. In general, researchers must decide what type of outcome measure is meaningful in the context of their study and what statistical method is best suited to that question.

Bayesian Sample Size Calculations

Adaptive Bayesian methods can offer several advantages over traditional frequentist study designs. These include:

- Shifting more patients to efficacious trial arms (or, conversely, shifting patients away from ineffective treatment arms)
- Requiring smaller overall study populations without loss of power and/or type I error
- Shortening the duration of the trial

In a simulation study of clinical trials for glioblastoma, researchers found that the use of Bayesian adaptive randomization resulted in 15% to 18% shorter trial duration, 30 fewer required patients overall without loss of statistical power, and more patients assigned to efficacious treatment arms (12 more out of a total sample of 140 in one specific scenario).

Source: Casewriters; Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale N.J.: L. Erlbaum Associates, 1988. Print.; Lorenzo Trippa, Eudocia Q. Lee, Patrick Y. Wen, Tracy T. Batchelor, Timothy Cloughesy, Giovanni Parmigiani, and Brian M. Alexander, "Bayesian Adaptive Randomized Trial Design for Patients with Recurrent Glioblastoma," *J Clin Oncol*, Sep 2012; 30(26): 3258-63.

Endnotes

- ¹ Susan Scutti, "Sen. John McCain has Brain Cancer, Aggressive Tumor Surgically Removed," *CNN*, July 20, 2017, <http://www.cnn.com/2017/07/19/health/gupta-mccain-glioblastoma/index.html>, accessed September 2017.
- ² American Association of Neurological Surgeons, "Glioblastoma Multiforme," <http://www.aans.org/Patients/Neurosurgical-Conditions-and-Treatments/Glioblastoma-Multiforme>, accessed May 2017.
- ³ American Cancer Society, "Survival Rates for Selected Brain and Spinal Cord Tumors," <https://www.cancer.org/cancer/brain-spinal-cord-tumors-adults/detection-diagnosis-staging/survival-rates.html>, accessed May 2017.
- ⁴ Lorenzo Trippa, Eudocia Q. Lee, Patrick Y. Wen, Tracy T. Batchelor, Timothy Cloughesy, Giovanni Parmigiani, and Brian M. Alexander, "Bayesian Adaptive Randomized Trial Design for Patients with Recurrent Glioblastoma," *J Clin Oncol*, Sep 2012; 30(26): 3258-63.
- ⁵ American Brain Tumor Association, "Glioblastoma and Malignant Astrocytoma," 2017, <http://www.abta.org/secure/glioblastoma-brochure.pdf>, accessed May 2017.
- ⁶ National Cancer Institute, "Cancer Stat Facts: Brain and Other Nervous System Cancer," <https://seer.cancer.gov/statfacts/html/brain.html>, accessed May 2017.
- ⁷ American Association of Neurological Surgeons, "Glioblastoma Multiforme."
- ⁸ American Brain Tumor Association, "Glioblastoma and Malignant Astrocytoma."
- ⁹ National Cancer Institute, "Cancer Stat Facts."
- ¹⁰ American Brain Tumor Association, "Glioblastoma and Malignant Astrocytoma."
- ¹¹ American Association of Neurological Surgeons, "Glioblastoma Multiforme."
- ¹² BM Alexander, PY Wen, L Trippa, DA Reardon, WK Yung, G Parmigiani, and DA Berry, "Biomarker-based Adaptive Trials for Patients with Glioblastoma – Lessons from I-SPY 2," *Neuro-Oncology* 2013; 15(8): 972-8.
- ¹³ Donald A. Berry, "The Brave New World of Clinical Cancer Research: Adaptive Biomarker-Driven Trials Integrating Clinical Practice with Clinical Research." *Molecular Oncology* 9 (2015): 951-9.
- ¹⁴ Donald A. Berry, "The Brave New World of Clinical Cancer Research."
- ¹⁵ Maria Kabisch, Christian Ruckes, Monika Seibert-Grafe, and Maria Blettner, "Randomized Controlled Trials." *Dtsch Arztebl Int.* 2011 Sep; 108(39): 663-8.
- ¹⁶ Maria Kabisch et al., "Randomized Controlled Trials."
- ¹⁷ Donald A. Berry, "The Brave New World of Clinical Cancer Research."
- ¹⁸ Jianchang Lin, Li-An Lin, and Serap Sankoh, "A General Overview of Adaptive Randomization Design for Clinical Trials," *J Biom Biostat* 2016; 7:2.
- ¹⁹ Jianchang Lin et al., "A General Overview of Adaptive Randomization Design for Clinical Trials."
- ²⁰ Stephan Ehrhardt, Lawrence J. Appel, and Curtis L. Meinert, "Trends in National Institutes of Health Funding for Clinical Trials Registered in ClinicalTrials.gov," *JAMA* 2015; 314(23): 2566-7.
- ²¹ John LaMattina, "Pharma Controls Clinical Trials of Their Drugs. Is This Hazardous to Your Health?" *Forbes*, October 2, 2013, <https://www.forbes.com/sites/johnlamattina/2013/10/02/pharma-controls-clinical-trials-of-their-drugs-is-this-hazardous-to-your-health/#727af2357b2e>, accessed June 2017.
- ²² Batelle and PhRMA, "Biopharmaceutical Industry-Sponsored Clinical Trials: Impact on State Economies," March 2015, <http://phrma-docs.phrma.org/sites/default/files/pdf/biopharmaceutical-industry-sponsored-clinical-trials-impact-on-state-economies.pdf>, accessed July 2017.

- ²³ Jason Monteleone, "Patient Recruitment: Clinical Research's 'White Whale'?" *Pivotal Financial Consulting, LLC*, December 12, 2016, <https://www.pivotalfinancialconsulting.com/single-post/2016/12/09/Patient-Recruitment-Clinical-Researchs-White-Whale>, accessed September 2017.
- ²⁴ U.S. Food & Drug Administration, "Step 3: Clinical Research," April 26, 2017, <https://www.fda.gov/forpatients/approvals/drugs/ucm405622.htm>, accessed May 2017.
- ²⁵ U.S. Food & Drug Administration, "Step 3: Clinical Research."
- ²⁶ California Biomedical Research Association, "Fact Sheet: New Drug Development Process," <https://www.ca-biomed.org/pdf/media-kit/fact-sheets/CBRADrugDevelop.pdf>, accessed May 2017.
- ²⁷ Batelle and PhRMA, "Biopharmaceutical Industry-Sponsored Clinical Trials: Impact on State Economies."
- ²⁸ U.S. Food & Drug Administration, "Drug Approval Process," <https://www.fda.gov/downloads/drugs/resourcesforyou/consumers/ucm284393.pdf>, accessed June 2017.
- ²⁹ Michael S. Kinch, Austin Haynesworth, Sarah L. Kinch, and Denton Hoyer, "An Overview of FDA-Approved New Molecular Entities: 1827-2013," *Drug Discovery Today* 19(8): 1033-9.
- ³⁰ Christopher P. Adams and Van V. Branter, "Estimating the Cost of New Drug Development: Is It Really \$802 Million?" *Health Aff* March 2006: Vol. 25, No. 2: 420-8.
- ³¹ Jason Millman, "Does It Really Cost \$2.6 Billion to Develop a New Drug?" *The Washington Post*, November 18, 2014, https://www.washingtonpost.com/news/wonk/wp/2014/11/18/does-it-really-cost-2-6-billion-to-develop-a-new-drug/?utm_term=.4759ee56e1b0, accessed June 2017.
- ³² Michael Hay et al., "Clinical Development Success Rates for Investigational Drugs," *Nature Biotechnology* January 2014; 32(1): 40-51.
- ³³ BM Alexander et al., "Biomarker-based Adaptive Trials for Patients with Glioblastoma – Lessons from I-SPY 2."
- ³⁴ Mohammed I. Zia, Lillian L. Siu, Greg R. Pond, and Eric X. Chen, "Comparison of Outcomes of Phase II Studies and Subsequent Randomized Control Studies Using Identical Chemotherapeutic Regimens," *J Clin Oncol* 2005, 23:6982-91.
- ³⁵ Mohammed I. Zia et al., "Comparison of Outcomes of Phase II Studies and Subsequent Randomized Control Studies Using Identical Chemotherapeutic Regimens."
- ³⁶ BM Alexander et al., "Biomarker-based Adaptive Trials for Patients with Glioblastoma – Lessons from I-SPY 2."
- ³⁷ Stuart A. Grossman, Xiaobu Ye, Steven Piantadosi, Serena Desideri, Lous B. Nabors, Myrna Rosenfeld, and Joy Fisher, "Survival of Patients with Newly Diagnosed Glioblastoma Treated with Radiation and Temozolomide on Research Studies in the United States," *Clin Cancer Res* 2010; 16(8): 2443-49.
- ³⁸ Peter Loftus, "FDA Approves Merck's Keytruda to Treat Cancers with Genetic Defects," *The Wall Street Journal*, May 23, 2017, <https://www.wsj.com/articles/fda-approves-mercks-keytruda-to-treat-cancers-with-genetic-defects-1495568988>, accessed August 2017.
- ³⁹ Peter Loftus, "FDA Approves Merck's Keytruda to Treat Cancers with Genetic Defects."
- ⁴⁰ Donald A. Berry, "Bayesian Clinical Trials." *Nat Rev Drug Discov*, 2006; 5(1): 27-36.
- ⁴¹ Donald A. Berry, "Bayesian Clinical Trials."
- ⁴² Swati Biswas, Diane D. Liu, J. Jack Lee, and Donald A. Berry, "Bayesian Clinical Trials at the University of Texas M.D. Anderson Cancer Center," *Clinical Trials* 2009; 6: 205-16.
- ⁴³ Donald A. Berry, "The Brave New World of Clinical Cancer Research."
- ⁴⁴ A.D. Stern, B.M. Alexander, and A. Chandra, "How Economics Can Shape Precision Medicines," *Science* March 17, 2017, 355(6330): 1131-3.
- ⁴⁵ Donald A. Berry, "The Brave New World of Clinical Cancer Research."

- 46 A.D. Stern, B.M. Alexander, and A. Chandra, "How Economics Can Shape Precision Medicines."
- 47 Swati Biswas et al., "Bayesian Clinical Trials at the University of Texas M.D. Anderson Cancer Center."
- 48 Donald A. Berry, "Bayesian Clinical Trials." *Nat Rev Drug Discov*, 2006; 5(1): 27-36.
- 49 Swati Biswas et al., "Bayesian Clinical Trials at the University of Texas M.D. Anderson Cancer Center."
- 50 Amelie Elsäßer et al., "Adaptive Clinical Trial Designs for European Marketing Authorization: A Survey of Scientific Advice Letters From the European Medicines Agency," *Trials* 2014 15(383).
- 51 U.S. Food and Drug Administration, "Adaptive Design Clinical Trials for Drugs and Biologics," February 2010, <https://www.fda.gov/downloads/drugs/guidances/ucm201790.pdf>, accessed June 2017.
- 52 U.S. Food and Drug Administration, "Adaptive Designs for Medical Device Clinical Studies," July 27, 2016, <https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm446729.pdf>, accessed June 2017.
- 53 Kelly Servick, "Congress and FDA Nominee Heap Love on 'Adaptive Trials,'" *Science*, April 7, 2017, <http://www.sciencemag.org/news/2017/04/congress-and-fda-nominee-heap-love-adaptive-trials>, accessed June 2017.
- 54 Senate Committee on Health, Education, Labor and Pensions, Senate Committee on Health, Education, Labor and Pension Hearing on the Nomination of Scott Gottlieb to be Commissioner of Food and Drugs, April 5, 2017, via ProQuest, accessed July 2017.
- 55 Scott M. Berry, Jason T. Connor, and Roger J. Lewis, "The Platform Trial: An Efficient Strategy for Evaluating Multiple Treatments," *JAMA* April 28, 2015; 33(16): 1619-20.
- 56 Scott M. Berry et al., "The Platform Trial."
- 57 Donald A. Berry, "The Brave New World of Clinical Cancer Research."
- 58 Shein-Chung Chow and Mark Chang, "Adaptive Design Methods in Clinical Trials – A Review," *Orphanet J Rare Dis* 2008; 3(11).
- 59 Shein-Chung Chow and Mark Chang, "Adaptive Design Methods in Clinical Trials – A Review."
- 60 Shein-Chung Chow and Mark Chang, "Adaptive Design Methods in Clinical Trials – A Review."
- 61 Shein-Chung Chow and Mark Chang, "Adaptive Design Methods in Clinical Trials – A Review."
- 62 Jack West Howard, "Novel Precision Medicine Trial Designs: Umbrellas and Baskets," *JAMA Oncology* March 2017 3(3): 423.
- 63 Donald A. Berry, "The Brave New World of Clinical Cancer Research."
- 64 AD Barker, CC Sigman, GH Kelloff, NM Hylton, DA Berry, and LG Esserman, "I-SPY 2: An Adaptive Breast Cancer Trial Design in the Setting of Neoadjuvant Chemotherapy," *Clin Pharmacol Ther* 2009 Jul; 86(1): 97-100.
- 65 AD Barker et al., "I-SPY 2: An Adaptive Breast Cancer Trial Design in the Setting of Neoadjuvant Chemotherapy."
- 66 I-SPY Trials, "Phase I: I-SPY Phase 1," <http://www.ispytrials.org/trials>, accessed June 2017.
- 67 Janet Woodcock and Lisa M. LaVange, "Master Protocols to Study Multiple Therapies, Multiple Diseases, or Both," *N Engl J Med*, 2017; 377: 62-70.
- 68 Penn Medicine News, "Two New Drugs Show Promise for Patients with Aggressive Breast Cancer," August 11, 2016, <https://www.pennmedicine.org/news/news-releases/2016/august/two-new-drugs-show-promise-for>, accessed August 2017.
- 69 Daron Acemoglu and Joshua Linn, "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry," *The Quarterly Journal of Economics*, August 2004; 1049-90.
- 70 Pierre Dubois, Olivier de Mouzon, Fiona Scott-Morton, and Paul Seabright, "Market Size and Pharmaceutical Innovation," *RAND Journal of Economics*, 2015; 46(4): 844-71.

- ⁷¹ BM Alexander et al., “Biomarker-based Adaptive Trials for Patients with Glioblastoma – Lessons from I-SPY 2.”
- ⁷² Brian M. Alexander, Evanthia Galanis, W.K. Alfred Yung, Karla V. Ballman, James M. Boyett, Timothy F. Cloughsey, John F. Degroot, Jason T. Huse, Bhupinder Mann, Warren Mason, Ingo K. Mellinghoff, Tom Mikkelsen, Paul S. Mischel, Brian P. O’Neill, Michael D. Prados, Jann N. Sarkaria, Abdul Tawab-Amiri, Lorenzo Trippa, Xiaobu Ye, Keith L. Ligon, Donald A. Berry, and Patrick Y. Wen, “Brain Malignancy Steering Committee Clinical Trials Planning Workshop: Report from the Targeted Therapies Working Group,” *Neruo Oncol* Feb 2015; 17(2): 180-8.
- ⁷³ National Biomarker Development Alliance, “Anna D. Barker, Ph.D.,” <http://nbdabiomarkers.org/bio/anna-d-barker-phd>, accessed September 2017.
- ⁷⁴ Anna D. Barker, “GBM AGILE – A Story of Convergence, Commitment, Collaboration and Compassion,” *International Brain Tumour Alliance*, June 18, 2016, <http://theibta.org/brain-tumour-magazine-201617/gbm-agile-a-story-of-convergence-commitment-collaboration-and-compassion/>, accessed July 2017.
- ⁷⁵ Brian M. Alexander, Sujuan Ba, Mitchel S. Berger, Donald A. Berry, Webster K. Cavenee, Susan M. Chang, Timothy F. Cloughesy, Tao Jiang, Mustafa Khasraw, Wenbin Li, Robert Mittman, George H. Poste, Patrick Y. Wen, W. K. Alfred Yung, and Anna D. Barker, “Adaptive Global Innovative Learning Environment for Glioblastoma: GBM AGILE,” *Clin Canc Res* (accepted for publication but not yet copyedited), <http://clincancerres.aacrjournals.org/content/clincanres/early/2017/08/16/1078-0432.CCR-17-0764.full.pdf>.
- ⁷⁶ Teresa Wang, Tej Azad, and Ritu Rajan, “The Emerging Influence of Digital Biomarkers on Healthcare,” *Rock Health*, 2017, <https://rockhealth.com/reports/the-emerging-influence-of-digital-biomarkers-on-healthcare/>, accessed September 2017.
- ⁷⁷ Frank R. Lichtenberg, Rebecca A. Pulk, Marc S. Williams, Eric Wright, “The Social Cost of Suboptimal Medication Use and the Value of Pharmacogenomic Information: Evidence from Geisinger,” September 2017, *Conference Presentation: Economic Dimensions of Personalized and Precision Medicine Conference (National Bureau of Economic Research)*.