Precision medicine is positioned to revolutionize the healthcare landscape, and the utilization of advanced clinical development technologies is key to supporting precision medicine trials. These technologies can harness a variety of disparate data sources at richer and greater volumes to build more accurate predictive models for different patient subsets. We discuss briefly the nature of precision medicine, the FDA’s initiatives encouraging the development of targeted therapies and the importance of the revolution in genomics for precision medicine. We describe the five Vs of big data and the pivotal role of big data in precision medicine. We present four clinical trial technologies that provide essential capabilities for precision medicine:

(1) Continuous collection, aggregation, and integration of myriad data sources; (2) Adaptive and iterative study design and execution as new information becomes available; (3) Maintenance and management of data throughout the study life cycle; and (4) Advanced analytics for research discovery. Precision medicine shows great promise for developing exploratory hypotheses from its broad variety of data and iterative approach, evaluating these hypotheses, and producing safer, more effective treatments for targeted patient subgroups with advantageous benefit-risk profiles. The complexities of precision medicine study designs require substantial planning, investment, and commitment from all stakeholders.
1. Introduction

Precision medicine is an approach to medicinal practice that tailors decisions to a patient’s individual characteristics: genetic, anatomical, physiological, environmental, and lifestyle. Precision medicine is defined by the National Academies of Science as “an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person” (National Academies of Science, 2017). Individually tailored healthcare is a rapidly developing, exciting alternative to the traditional practices that are often described as one-size-fits-all, based on the best course of prevention or treatment for the general population. In this approach, if an initial treatment fails, the patient moves on to the next standard treatment, again based on the population. The challenge of this method of practicing medicine have long been recognized.

Clinicians and patients today have access to a rapidly expanding range of data. Drug regulatory authorities and the practice of medicine itself are adapting to this new information era, which holds promise both to further our understanding of disease etiology and to deliver better patient care and optimal outcomes (Heckman-Stoddard, 2014; Woodcock, 2017).

Torkamani et al. (2017) recently expanded on the concept of precision medicine by defining high-definition medicine as “the data-driven practice of medicine through the utilization of highly detailed, longitudinal, and multi-parametric measures of the determinants of health to modify disease risk factors, detect disease processes early, drive precise and dynamically adjusted interventions, and determine preventative and therapeutic intervention efficacy from real-world outcomes.” Establishing a personal baseline through frequent assessment of the determinants of health can enable early detection and intervention, sometimes before diseases manifest clinically.

The objective of this paper is to describe the current state of affairs for sponsors who are considering a precision medicine approach and seeking to create a development path that will maximize the chance of success.

1.1. Precision Medicine and the FDA

The U.S. Food and Drug Administration (FDA) has embraced the possibilities offered by precision medicine. This is reflected in a recent New England Journal of Medicine article by Janet Woodcock, director of FDA’s Center for Drug Evaluation and Research (CDER), and Lisa LaVange, director of the Office of Biostatistics at CDER (Woodcock, 2017) that discusses the need to capitalize on similarities among trials (e.g., shared control group and recruiting efforts) to gain efficiencies in drug development using umbrella, basket, or platform trials and cites I-SPY2 trial and Lung-Map as examples of such innovation.

The FDA recently announced the Precision Medicine Initiative (PMI), with the overarching goals of facilitating the identification of genetic drivers of disease and developing better diagnostic tools and treatments (FDA, 2016). The FDA has issued two PMI draft guidances and established the precisionFDA portal for oversight, analysis, and validation of methodology related to next-generation sequencing (NGS), also known as high-throughput sequencing. NGS tests are used to identify the genomic differences in individuals (FDA, 2016) that are the basis of the genetic component of precision medicine.

The FDA has fostered numerous initiatives that encourage drug sponsors to develop targeted therapies and has provided guidance on the regulatory pathway for such medicines. FDA’s Enrichment Guidance (FDA, 2012) describes approaches to studying therapeutics when the efficacy or safety of that therapeutic is expected to vary across subgroups. The FDA guidance on adaptive designs provides advice for incorporating analyses of accumulating data at prospectively planned time points (FDA, 2010). Adaptations such as modifications in the treatment randomization ratios, dropping or adding an arm, or revising a patient’s treatment based on their inadequate response to an initial treatment are all examples of how adaptive designs can fuel precision medicine and are discussed in the guidance. The Co-development Guidance provides advice for obtaining contemporaneous marketing authorization for a therapeutic product and its corresponding companion diagnostic (FDA, 2016). And, expected soon is the Complementary Diagnostics Guidance.

This guidance is expected to lay out a framework that allows for a diagnostic device that aids in the benefit-risk assessment of the therapeutic product (but it is not essential as with the companion diagnostic).

The FDA has approved an increasing number of medicines that are indicated for defined subsets of patients, including the first ever-approval of a drug [Keytruda (pembrolizumab)] to treat patients whose cancers have a specific genetic feature, or biomarker, rather than being based on the location of origin of the cancer. Importantly, these precision medicine approvals are not limited to oncology indications. For instance, the FDA has approved two products for treatment of severe asthma in patients with an eosinophilic phenotype [Cinqair (reslizumab), Nucala (mepolizumab)]. In March 2017, FDA concurrently approved a drug [Zejula (niraparib)] and a complementary diagnostic test [BRACAnalysis CDx] to identify patients with germline BRCA mutations, the subset of patients expected to benefit most from Zejula. A complementary diagnostic test is defined as one that aids in the benefit-risk decision(s) related to a particular therapeutic product but is not essential to the safe and effective use of the product, as is the case with a companion diagnostic. Draft guidance on the topic of therapeutic products developed with a complementary diagnostic test is expected from the FDA soon (Mansfield, 2015). Overall, the FDA continues to solidify a workable regulatory platform that encourages innovation while maintaining efficacy and safety as priorities.
1.2. Genomic Data and the Current State of Precision Medicine

The ongoing revolution in genomics and molecular biology continues to improve our fundamental understanding of human biology and disease; at the same time, it presents new opportunities for precision medicine. For example, not long ago, breast cancer was largely defined as a single disease. Today, based on molecular subtyping, breast cancer is recognized to consist of several distinct molecular subtypes (Bartlett, 2017). The biomarkers that distinguish these molecular subtypes or influence the risk of cancer development have been translated into targeted clinical practice (Marchina, 2010; Godet, 2017; Slamon, 2001). Similarly, the clinical phenotype of depression has long been known to encompass several subtypes (e.g., major depressive disorder, bipolar disorder, and atypical depression) and to show high variability regarding treatment response. This has often made the diagnosis and treatment of depression challenging for clinicians. However, exciting new biomarker data may soon aid clinical practice by providing-molecular insights to guide accurate diagnosis and prediction of treatment response (Buzuk, 2016; McIntyre, 2016; Strawbridge, 2017; Yue, 2016).

Increasingly, precision medicine is transitioning from a specialized treatment approach to an aspect of standard clinical practice. As of September 2017, approximately 179 studies (~50% actively recruiting) were listed in clinicaltrials.gov as search returns for the term “precision medicine,” with approximately 21% of these trials being funded by industry. This is a gross underestimate of the actual number of trials that can be deemed precision medicine trials, because most of these trials are not explicitly labeled as such in clinicaltrials.gov. Based on combinations of keywords related to biomarkers, next generation sequencing and synonymous terms, we can roughly suggest that this number is an order of magnitude greater.

As precision medicine continues to propagate into clinical practice, it is increasingly important for clinicians to be trained to understand precision medicine and to use it effectively for the benefit of their patients. The National Human Genome Research Institute (a research institute of the National Institutes of Health (NIH)) recently awarded several training grants that will instruct medical doctors in applying computational tools to disease-gene discovery, patient genome interpretation, and big data management in research and clinical settings (National Human Genome Research Institute, 2016).

Drug developers may view precision medicine as a daunting and uncertain task, and utilizing large datasets (i.e., big data) effectively for precision medicine requires careful planning, investment, and discussion among stakeholders. Establishing the appropriate infrastructure is a crucial step in dealing with huge volumes of information and the complexities of adaptive precision medicine study designs.
To predict outcomes accurately in a patient subgroup, precision medicine demands large volumes of data from often disparate data streams, some of which are generated continuously. One challenge for drug developers is how to seamlessly ingest, aggregate, integrate, and analyze these datasets in an easy-to-use workflow that can help uncover meaningful relationships resulting in improved patient outcomes. Precision medicine is an exploratory and iterative journey, so the data platform of choice must be flexible enough to adapt to emerging data and to allow investigators to redefine – in real time – patient clusters into more precise subsets and to redesign studies to enable refined hypothesis testing. A useful starting point is to adopt a flexible, integrated, and intelligent clinical development platform as the engine that will drive the precision medicine program. These are crucial features in an optimized precision medicine approach.

The remainder of this paper discusses the relevance of big data to precision medicine trials and some key clinical development technologies that can facilitate the development of targeted therapies.

The ability to redefine patient clusters and adapt study designs in real time are key attributes of a platform to support a precision medicine approach.

As drugs become more precise and trials become more complex, it will be increasingly important to exploit optimized clinical trial technologies. The platform should seamlessly ingest, aggregate, integrate, and facilitate analysis of disparate data types so that the ideal target patient population can be identified quickly and reliably.
2. Big Data and the Path to Precision Medicine

A precision medicine approach to drug development relies on large datasets to cluster patients into increasingly more precise subsets. This requires that investigators measure patients on numerous attributes of many different types, ranging from genotype to phenotype and beyond (Beckmann, 2016; Dinov, 2016; Dumitrescu, 2017). Simply ingesting and managing data of such scope and complexity are formidable tasks, and conceiving, executing, and interpreting appropriate analyses are even more formidable. However, we must meet these challenges in order to achieve the tremendous gains that precision medicine promises over one-size-fits-all population-based medicine.

The precision medicine approach generates a unique and comprehensive patient profile by integrating a wide range of medical data. Taken in the aggregate, these data encompass enormous volume, velocity, variety, veracity, and value. We can describe these attributes – the five Vs of big data – briefly as:

- **Volume**: Enormous amounts of data are generated by many data streams of different types, including over 3 billion nucleotides in the human genome; hundreds or thousands of laboratory test results and vital signs; wearable sensor readings of dozens of kinds; environmental data on dozens or hundreds of relevant biological, chemical, and meteorological features; medical imaging data containing many thousands or millions of pixels per image; and more.

- **Velocity**: Many data streams produce new data at rates of dozens or hundreds of points per second (e.g., sensor data, environmental data, imaging data).

- **Variety**: The information collected for a patient includes structured and unstructured data from vital signs, laboratory tests, genomic sequencing, patient behaviors, social media, wearable sensor readings, point-of-care notes, photos, medical images, and more.

- **Veracity**: It is critical for precision medicine trials to have high-quality data. Missing and inaccurate observations can rarely be completely eliminated, but every effort must be made to minimize their occurrence. Poor data quality will produce errors in data interpretation and the analyses required for precision medicine.

- **Value**: The impressive scope of big data is not an end in itself, but a means to the end of valuable results: better therapeutic and preventative treatments. Precision medicine must manage the volume, velocity, and variety of big data and maximize the veracity of the data; however, to be successful, precision medicine must extract value from the data with analytical algorithms whose results yield improved treatments for patients.

The utility of genetic information to drive precision medicine has been clearly demonstrated (e.g., trastuzumab (Herceptin) for women with HER-2 positive breast cancer). Other important data streams also offer vast possibilities for additional lines of investigation. Electronic health records (EHRs) are an example of big data that can help inform a precision medicine approach. A recent article by Dumitrescu et al. (2017) concluded that current and new EHR technology will provide international standards for applications that facilitate precision medicine using the complex healthcare information in health, social, economic, behavioral, and environmental data. This article reported the results of a genome-wide association study for resistant hypertension linked to EHRs and genotypes. The authors demonstrated the utility of DNA biobanks linked to EHRs for classifying resistant hypertension cases and controls and for performing large-scale genome-wide association studies. Evans (2016) utilized an algorithm based on a combination of billing codes, laboratory values, text queries, and prescription records. A study by Szlosek (2016) reported the use of machine learning and natural processing algorithms to automate the evaluation of clinical decision support in EHR systems. The authors showed that a machine learning algorithm can be used to identify abnormal head CT scans in free-text health care data with a high degree of accuracy. These examples demonstrate that using big data to support a precision medicine approach has moved beyond a conceptual framework. These and other well-documented studies showcase how big data technology can enable precision medicine. However, it is our experience that some drug developers are not prepared to handle the complexities inherent in the five Vs of big data, which can quickly overwhelm the clinical development platform and the supporting business processes not designed for them.
In precision medicine as well as traditional medicine, data quality is a wide-ranging, serious issue that poses a grave threat to the validity of clinical research. Poor planning often results in poorly standardized data, missing or incomplete clinical datasets, incompatible conceptualizations of data types and elements, and infrastructure that is inadequate for processing robust healthcare data (Anderson, 2014; Field & Sansone, 2016; Mead, 2006; Richesson, 2007).

The following section highlights clinical trial technologies that make precision medicine possible, not only by enabling the efficient collection, aggregation, and integration of data, but by empowering the exploratory and iterative processes that drive the precision medicine approach via powerful analytics.

3. Clinical Trial Technologies to Facilitate Precision Medicine

Advanced clinical trial technologies allow investigators to seamlessly collect and integrate massive amounts of data from sources across the clinical care continuum. An adaptable system can leverage both historical datasets and in-study datasets to update and optimize models that allow for re-running the cohort selection, risk stratification, and subsequent intervention assignment. Further, advanced predictive analytics enable the development of new treatment models that can adapt to new information as it is uncovered, with subsequent forecasts and trials becoming increasingly accurate as the volume of data expands.

As discussed earlier, an important first step to achieving a precision medicine approach is to utilize a clinical technology platform that ingests a wide variety of phenotypic data, integrates the data into standardized data models, and unifies these data models with analytic technologies. This is a key determinant of an optimized precision medicine approach. Huge volumes of data, poorly standardized data models, or integration challenges between data models and analytical technologies can quickly overwhelm ill-equipped data centers and business processes, and without an adequate infrastructure in place, the entire platform can collapse, leading to lost time, money, and even valuable data.

The platform should provide these capabilities, which we list and then discuss in more detail:

- Continuous collection, aggregation, and integration of myriad data sources.
- Adaptive and iterative study design and execution as new information is made available.
- Maintenance and management of data throughout the study life cycle.
- Advanced analytics for research discovery during the study and post-study.

3.1. Continuous Collection, Aggregation, and Integration of Myriad Data Sources

To benefit fully from insights available with a big data approach, massive amounts of data must be collected, aggregated, and integrated, either in real time (e.g., sensor data) or from archives (e.g., unstructured data such as point-of-care notes, histological images, and photos). By capturing a wider variety of non-traditional data derived as close to the source as possible, big data offers amazing opportunities for precision medicine. Below are brief descriptions of the types of data that are integral to informing a precision medicine approach.

Laboratory data and vital signs continue to serve as crucial data points. These data provide temporally important insights; for example, blood work can evaluate the status of a biomarker to aid in diagnosis or prognosis of a disease.

With the advent of NGS, genetic analyses can have increasing impact on precision medicine; NGS can simultaneously target hundreds of genes that are known to be important for a disease (Roychowdhury, 2016). The genetic testing firm Concert Genetics estimates that there are more than 65,000 genetic testing products available on the market, with 10 new products introduced per day. As additional genetic data are verified, this will allow for the definition of more precise stratification of patient subsets and better assessments of interactions between genotypes and clinical phenotypes derived from biobank data (Azimi, 2016). New high-quality genome assemblies are increasingly offering ethnicity-specific reference sequences that can be used to improve precision medicine studies in certain populations, such as Asians (e.g., Seo, 2016; Qiao, 2016; Shi, 2016).

Imaging data are a medically rich source that can also be a resource-intense burden. Much of the available technology is unable to analyze these data effectively without careful deep learning or neural network techniques (Dinov, 2016; Geraci, 2017). Continual advances in this realm are generating new ways to use technology to evaluate and even interpret images. In histology, for instance, much of the routine work of diagnosis follows algorithmic decision trees. Recent studies have reported success in using deep learning to increase the accuracy and efficiency of histopathological diagnoses (Djuric, 2017; Litjens, 2016). In another example, deep convolutional neural networks were shown to have potential for general and highly variable tasks across many fine-grained object categories: investigators demonstrated artificial intelligence (AI) to be capable of classifying skin cancer with a level of competence comparable to dermatologists (Esteva, 2017).

Digital biomarkers represent a valuable stream of citizen-generated data from sensors embedded in a variety of devices (e.g., wearables, implantables, and smartphones). These digital biomarkers can play a vital role
in helping to identify more precise subsets of patients, based on their ability to provide longitudinal data in a real-world setting (as opposed to a single snapshot in time in a doctor’s office). More advanced intuitive designs in microelectromechanical systems are providing researchers with abundant data to aid the precision medicine efforts for a variety of phenotypes of interest (Redmond, 2014). In addition to health-related data, a variety of behavioral and social data can also be used to explore social networks and social determinants of health for patient outcomes.

Prior to initiating a precision medicine trial, sponsors must develop their study designs based on the types of data that will be collected. Sponsors must also be prepared to address data quality deficiencies that result from aggregating disparate data sources with varying degrees of structure, and they must be able to integrate data quality assessment functionality to ensure the ongoing validity and reproducibility of findings. Developments in AI are expected to make data collection and integration even more efficient in the near future. For example, AI is actively being used in the form of chatbots for patient triage (e.g., Microsoft’s Health bot), and this technology is also being explored to support the burdensome process of creating clinical notes and phenotyping disease (Deliberato, 2017; Geraci, 2017). As data streams become increasingly structured and streamed in real time, the precision medicine approach will become even more efficient and accelerated.

3.2. Adaptive and Iterative Study Design and Execution as New Information is Made Available

Precision medicine studies must offer investigators the capability to adapt study designs throughout the life cycle of the study, based on newly added information or newly discovered relationships in the data. The precision medicine approach is initially exploratory, and it requires iteration to redefine patient clusters into more precise subsets and to redesign studies to enable refined hypothesis testing.

Investigators must find the right balance of the number of initial hypotheses that will be tested (i.e., exploratory analyses), since too many hypotheses can be viewed as unfavorable and lacking a defined research objective. Likewise, too many post hoc analyses constitute data dredging. While, understanding causal relationships in data to surface actionable insights that are truly governing improvements can have a significant impact. There is a need to use appropriate methods, avoiding model misspecifications, and careful selection of data to avoid issues of endogeneity, which can manifest from rate dependence or state dependence.

Methodological innovations in trial design include the ability to evaluate multiple treatments in more than one patient type or disease within the same overall trial structure (Woodcock, 2017). The key to this approach is to generate a master protocol, a single overarching protocol designed to answer multiple questions (e.g., studies involving multiple drugs and/or diseases) and to enable the use of “innovative statistical approaches to study designs and analysis, enabling a broader set of objectives to be met more effectively than would be possible in independent trials” (Woodcock, 2017). The master protocol can accommodate different clinical trial designs, including umbrella, basket, and platform trials (Woodcock, 2017).

Key tasks in designing a precision medicine study:

- Generate the protocol — establish solid scientific foundation for predictive precision medicine trial via pre-protocol literature research to derive several initial hypotheses to be tested (may result in generating a new master protocol or capitalizing on existing trial infrastructures).
- Plan adequately for the complex trial designs and real-time decision making that is the foundation of precision medicine trials, including a basis for decisions about treatments to pursue and those to discontinue.
- Consider a wide array of data types and depth to be measured, based on the range of relevant factors and outcomes.
- Construct the infrastructure required to handle data flow, data quality, data sharing, trial logistics, and real-time ability to analyze trial data and redesign trials according to new information.
- Evaluate innovative methods for analyses and recognize the care required to interpret results.

After the study objectives and measurable outcomes have been finalized in the protocol, downstream analyses may leverage some of these clinical trial designs:

- Longitudinal cohort studies: Precision medicine cohorts can be created to evaluate the relationships between genotype and phenotype over time, such as the NIH sponsored program All of Us (Collins, 2015; National Institutes of Health, 2017).
Basket trials: Enable a larger recruitment pool, since this design tests whether a drug is effective in patients characterized by a single genetic alteration and whether patients may have tumors of different histologic types and tissue of origin (Simon, 2017; Redig, 2015; American Society of Clinical Oncology, 2017). For instance, the National Cancer Institute Molecular Analysis for Therapy Choice (NCI-Match) is a Phase II clinical trial in which patients who share a common genetic mutation for a given cancer are sorted into “baskets,” or treatment arms, regardless of cancer type.

Umbrella trials: Assign patients to a particular treatment arm based on their cancer type and genetic marker. This design can have many treatment arms. The Lung Master Protocol (Lung-MAP) study aims to rapidly identify drug therapies for particular cancer types, making it a useful design for cancers with wide genetic heterogeneity (Woodcock, 2017).

N-of-1 studies: In these trials, which are often used in evaluation of musculoskeletal or pulmonary conditions, a patient serves as both control and experimental treatment. Although seemingly counterintuitive to the goal of precision medicine, which typically harnesses large amounts of data, this design is especially useful for low-frequency, rare conditions (Schork, 2015; Duan, 2013; Lillie, 2011).

Adaptive trials: Investigators can adapt study goals based on emerging data, thus provoking a revision of hypotheses to be tested (Heckman-Stoddard, 2014). Treatment arms can be altered or biomarker strategies adjusted depending on drug efficacy data (Woodcock, 2017). This paradigm is potentially conducive to more rapid drug approvals for precision medicine (Printz, 2013).

Figure: Schematic highlighting the flow of umbrella and basket trials (Redrawn from Woodcock, 2017).
The precision medicine approach requires rigorous pre-trial discussion between sponsors, clinical trial management firms, and the governing regulatory agency to ensure alignment with the protocol; therapeutic approaches that will be tested to improve patient outcomes; and timing of regulatory submission filing (Woodcock, 2017; Schwaederle, 2015).

"The use of a single system for clinical data management will enable shorter start-up times as the protocol is expanded to incorporate new investigations. The use of a single central randomization system facilitates the addition of new therapies with minimal disruption. Real-time access to the genomic, proteomic, pathological, and imaging data streams is requisite for the adaptive features of I-SPY 2 [Neoadjuvant and Personalized Adaptive Novel Agents to Treat Breast Cancer]."

(Woodcock, 2017)

Precision medicine trials necessitate a steadfast approach with considerable advance planning, because the protocols and trial designs can be quite complex.

3.3. Maintenance and Management of Data Throughout the Study Life Cycle
The complexities underlying precision medicine trials require substantial planning for data maintenance and management over the entire life cycle of the study. The highly adaptive nature of precision medicine trials demands that high-quality data be maintained throughout the trial.

Promoting high-quality data relies on stringent, consistently applied definitions of data quality as well as quality assurance and control procedures to ensure the validity of the incoming data and reproducibility of results. Data standards should be maintained to assess quality on a variety of dimensions (Kahn, 2016; Weiskopf, 2013). The breadth of information generated in precision medicine often warrants the development of procedures and statistical models and algorithms to ensure rapid and uninterrupted data flow (Woodcock, 2017; Garraway, 2013).

The infrastructure for logistical efficiencies that establish optimal data quality is a necessary component for the success of a precision medicine trial. Engaging in early discussions with experts in clinical trial technologies can help establish the trajectory for success.

3.4. Advanced Analytics for Research Discovery During the Study and Post-study
Precision medicine trials rely heavily on clinicians and researchers to consider emerging data in new ways and to develop innovative strategies to fully leverage the massive body of data generated.
“The collection and processing of biomarker samples is extremely important to the integrity and validity of the genetic data used in a precision medicine trial. Optimized clinical trial technologies use a combination of risk indicators and advanced pattern detection, to identify sites that may be lacking in the collection and processing of samples according to defined protocol procedures. The faster any quality issue with biomarker samples can be identified, the faster mitigation strategies can be deployed.”

Both during the study and following study close-out, advanced analytics can help uncover newly observed patient behaviors, biomarkers of response, new subgroups of responsive patients, new diagnostic criteria or risk factors, novel treatment paradigms, etc. This process can result in revised or even completely novel studies with increasingly focused hypotheses to be tested.

Post hoc or exploratory subgroup examination has traditionally been criticized for several well-documented reasons, including selection bias. However, developments in advanced analytics — including machine learning, data mining, and statistical methods — emphasize critical statistical principles and the importance of performing multiplicity adjustments to adjust for selection bias inherent in subgroup search. Some of these techniques also treat subgroup investigation as a special case of model selection (Lipkovich, 2017; Benjamini, 2005). Such advances are favorable to the type of post-study analyses required to generate novel insights to drive equally novel precision medicine approaches.

Figure: The utility of biomarkers in post-study analyses (Redrawn from Boyapati, 2016).
In planning a precision medicine trial, proper consideration must be given to the clinical development software platform that will be used for post-study analytics. The software should be user-friendly and integrated with the existing trial infrastructure, to provide for seamless workflows and to avoid the hefty costs and extra work required to integrate disparate solutions and adjusting to the various environments.

4. Summary

Precision medicine is positioned to revolutionize the healthcare landscape, and the utilization of advanced clinical development technologies is key to supporting precision medicine trials. These technologies can harness a variety of disparate data sources at richer and greater volumes to build more accurate predictive models for different patient subsets.

As the number of data streams increases across the continuum of clinical care, the promise is that precision medicines will become increasingly targeted. However, the complexities associated with evolving study designs are substantial, and the precision medicine approach demands solid planning, investment, and commitment from all stakeholders.

An integrative platform should be designed to house the clinical data; streamline standardized workflow in data collection, analysis, and iterative hypothesis testing; and provide a satisfactory clinical user experience. Our need to evaluate exploratory hypothesis testing requires the thoughtful construction of innovative clinical trials that are adaptive to data generated in real time. A platform that allows an iterative research process increases the likelihood of success for sponsors embarking on a precision medicine approach.

The promise of safer, more effective treatments through targeting of specific patient subgroups with the most advantageous benefit-risk profile is what propels researchers on the precision medicine journey.

Machine learning algorithms have invigorated post-study analyses by minimizing biases while allowing for novel insights that can be subsequently tested in a prospective trial.”

References


e/ucm2017930.pdf.


cements/materials/scienceboardtothefoodanddrugadministration/ucm315006.pdf.


Steven Schwager, PhD, is a Research Principal in the Data Science group at Medidata Solutions and a Professor Emeritus of Statistics and Biological Statistics at Cornell University. He joined the faculty at Cornell after earning his PhD in statistics at Yale University. His research includes statistical methodology and theory and collaborative projects in a wide range of biological and physical sciences and other fields. He works at Medidata on developing statistical and analytical methods for improving the design, execution, and analysis of clinical trials, including Clinical Trial Genomics, operational analytics, and synthetic control arms.

Jeffery Wiser, recently joined Medidata as Genomics Portfolio Director to grow the new Clinical Trial Genomics (CTG) product. Jeff has spent his 20 year career leading the development of life sciences software with a focus on the integration of clinical and ‘omics’ data. For the last 10 years, he has served as the Program Manager or Co-Investigator on multiple National Institute of Health funded projects, most notably the ImmPort project, a public/private partnership with UCSF and Stanford University. ImmPort serves as a repository for the collection, QC, curation and sharing of clinical and mechanistic assay results generated by researchers funded by the National Institute of Allergy and Infectious Diseases (NIAID), and was recently designated a Nature Scientific Data recommended repository for immunology and cytometry. Previously, Jeff served for 9 years in management roles at Gene Logic Inc. building clinical data capture, biobanking, LIMS, and gene expression analysis software for what was at the time the largest microarray gene expression database in the world. He holds a MSE from the University of Pennsylvania in Bioengineering and a BS from Cornell University in Mechanical Engineering.

Ruthie Davi, PhD, has worked in the area of pharmaceutical product development, and specifically the FDA, over 20 years. In her current position as a Director in Data Science at Medidata Solutions, she is part of a team focused on providing centralized systematic monitoring and other novel clinical trial tools. Ruthie has held numerous positions within the FDA, including Statistical Reviewer, Team Leader, and Deputy Director, where she implemented innovative statistical methods and made recommendations for clinical trial designs tailored to the regulatory setting. She has conducted the statistical review, represented the Agency at Advisory Committee meetings, and provided recommendations regarding FDA marketing approval for numerous New Drug and Biologic License Applications. With a particular interest in pediatric clinical trials, she was an active member of FDA’s Pediatric Review Committee. Ruthie holds a Ph.D. in biostatistics from George Washington University.