

# Using Data and Analytics in Clinical Development:

An Industry Perspective

AN eBook FROM MEDIDATA DRAWN FROM  
AN ONLINE SURVEY OF CLINICAL RESEARCH COMPANIES  
FIELDED IN CONJUNCTION WITH *PHARMAVOICE*



## Executive Summary

---

Together with *PharmaVoice*, Medidata Solutions sponsored a survey of executives within the life sciences industry to better understand how they are using data and analytics in clinical research. The online questionnaire was fielded in April and May of 2017 and garnered responses from 189 participants.

As clinical trials become more complex, how are stakeholders using data and analytics to improve their trials? The survey provides an opportunity to hear directly from those on the frontline of clinical development.

Respondents as a whole concern themselves most with the need to reduce trial timelines, followed by developing more effective drugs and devices. The emphasis placed on these two goals varies slightly for C-suite executives versus those in the next layer of management.

The need to reduce trial costs was mentioned more often within small and mid-sized companies than in the largest companies.

Respondents list several hurdles to achieving their clinical development objectives, chief among them improving patient recruitment, managing data quality, and strengthening patient retention.

Technology has enabled the proliferation of trial data from new sources, and it will also be the backbone of the solution. New methods of cleaning and monitoring trial data will be system-driven, iterative, and adaptive.

Patient retention is another significant issue, as the loss of enrolled patients threatens the validity of trials. It is also a waste given the effort and resources that are extended to recruit patients in the first place.

Almost all respondents rely on data and analytics to achieve their development goals. More companies rely on in-house data than not, but the use of analytical capabilities is fairly split between in-house and external resources.

Respondents have a need for a wide range of data types, but are particularly interested in real-world data (RWE) and historical trial data. Similarly, survey participants recognized value in various analytical techniques and methodologies. The industry would seem to be in the early stages of applying advanced analytics to increase the efficiency of development efforts—a fact that speaks to the promise of improvements to come.

We foresee that companies will be challenged to implement the latest technologies and will need to apply best practices in change management to be successful.

## Key Findings

- Reducing trial timelines is the highest priority for all companies of all sizes
- Minimizing the risks affecting regulatory approval is the top priority for nearly 1/3 of C-suite executives
- Small companies are more concerned about reducing trial costs than larger companies
- Companies' top three challenges: patient recruitment, data quality, patient retention
- 90% of respondents use data and analytics solutions
- Just over half of respondents rely on in-house data for clinical development insight
- The most sought-after data = real world evidence and historical trial data
- Most useful analytical capabilities = identifying patterns and anomalies, predictive analytics, and trial design optimization

## Respondent Demographics

Of the 189 survey respondents, 17.5% are in C-suite positions and 29.2% are in vice president and director-level positions, including leaders of therapeutic areas (See Figure 1). Nearly half (47.2%) of the respondents are in “Other” job roles, and these include such diverse jobs as: clinical research associate, consultant, analyst, clinical research supervisor, clinical trial manager, process analyst, medical editor, portfolio management, principal scientist, program manager, regulatory affairs, clinical project management, clinical operations, procurement, study placement, data management, laboratory, and study coordinator.

The survey drew participants from various types of companies for an all-around view of the clinical research space. Half of the respondents work in pharmaceutical companies, 16.9% for contract research organizations (CROs), 14% for biotech companies, and 6.6% for medical device manufacturers (Figure 2).

Figure 1: Respondents by Job Title

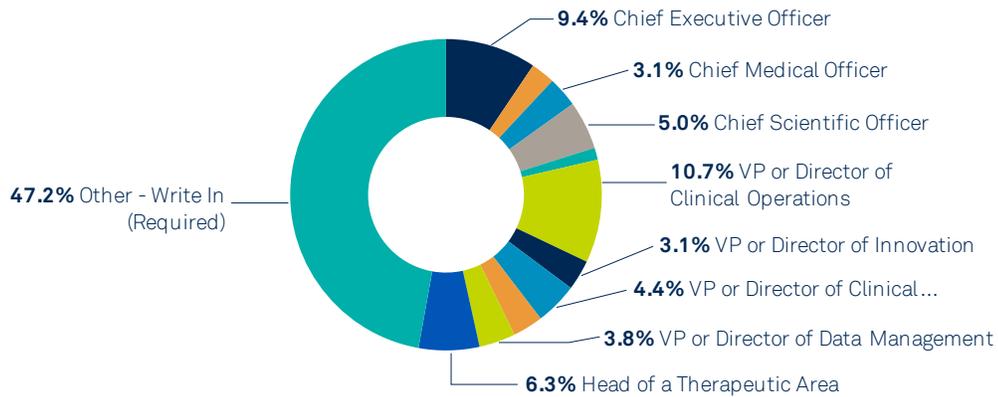
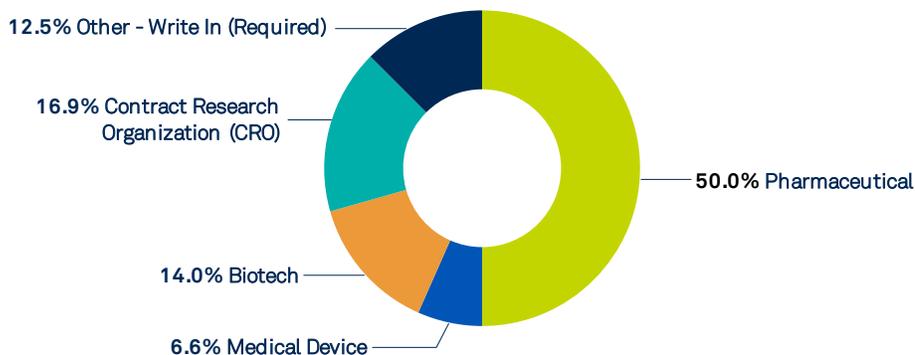


Figure 2: Respondents by Company Type



Similarly, participants represent companies of all sizes. A little over 40% work in small companies with revenues under \$250 million, another 12.5% are in companies with revenues between \$250 million and \$500 million, around 10% work in companies with revenues of \$500 million to \$1 billion, and roughly 40% are in companies of over \$1 billion in revenue (Figure 3).

The vast majority (77.6%) of respondents work in North America, while 5.3% are in Asia, 4.5% in the EU or Middle East, 1.3% in Japan, and 0.7% in Africa (Figure 4).

Figure 3: Respondents by Company Size

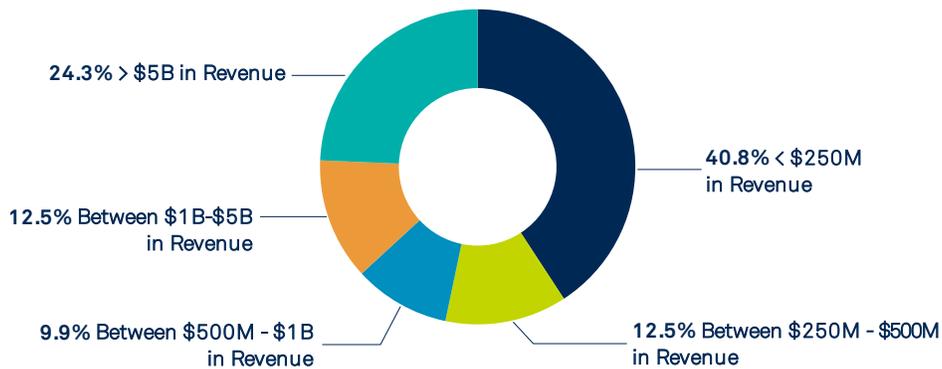
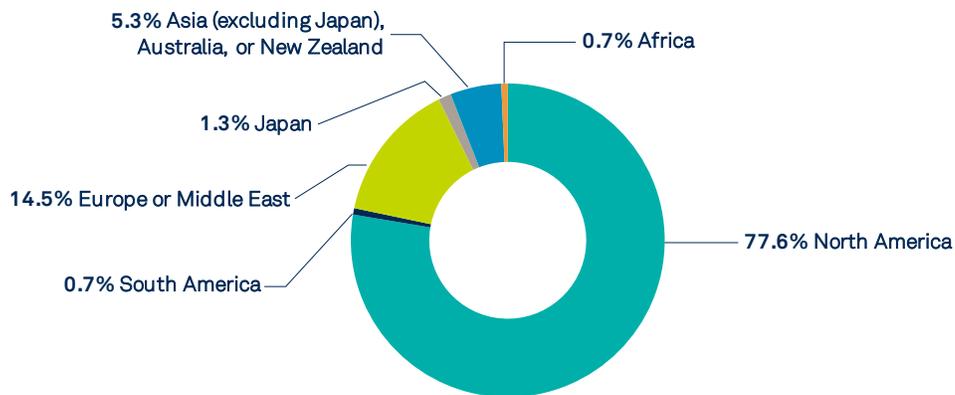


Figure 4: Respondents by Region



## What Is Your Organization’s Top Clinical Development Priority?

For respondents as a whole, the most pressing concern is reducing trial timelines leading up to regulatory submission (32%), followed by developing more effective drugs/devices (18.1%). Most respondents appear to be confident in the scientific aspects of trials, including their ability to make go/no-go decisions on drug candidates and to manage operational trial risks (Figure 5).

Respondents’ priorities are, naturally, closely tied to their roles/organizational levels (Figure 6). Executives in the C-suite are most focused on minimizing the risks affecting regulatory approval (30%), on reducing trial timelines (28%), and on achieving the company’s overarching mission of developing more effective drugs/devices (19%).

Respondents at the vice president and director levels are most concerned about how they can execute the mission efficiently. They rank reducing trial timelines (33%) as their highest priority, followed distantly by minimizing the risks that affect regulatory approval (18%), and reducing the cost of trials (16%).

Figure 5: Strategic Priorities for Respondents as a Whole

Of the options below, which is the highest strategic priority for your organization over the next 1-3 areas in the area of clinical development/clinical trials?

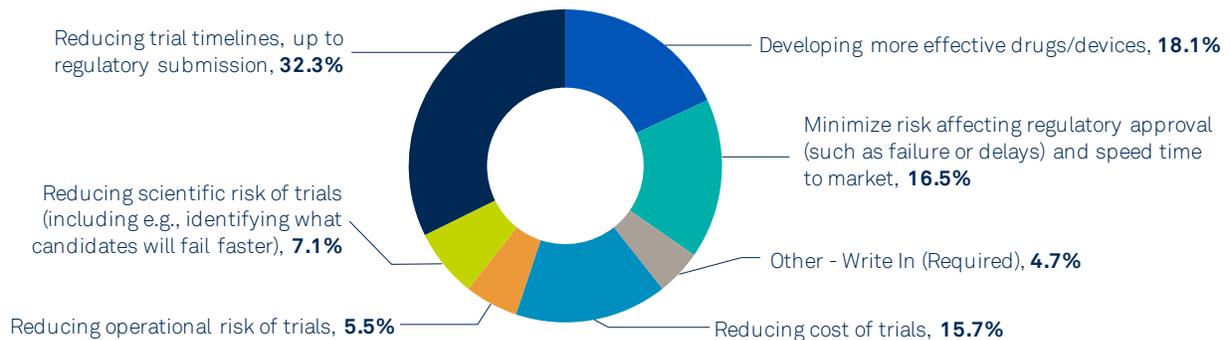
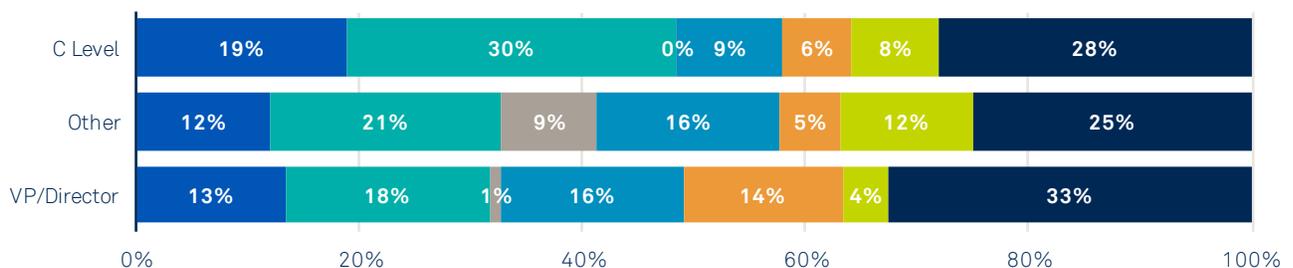


Figure 6: Strategic Priorities by Job Level



## Strategic Priorities by Company Size

Interestingly, the ranking of strategic priorities does not change by type of company, although it changes slightly by size of the company (Figure 7).

Companies of all sizes named reducing trial timelines as their top priority with 76% of larger companies (\$1B and up), 26% of medium-sized companies, and 22% of the smallest companies selecting with this option. For medium-sized companies, this concern was tied with reducing regulatory risk in importance.

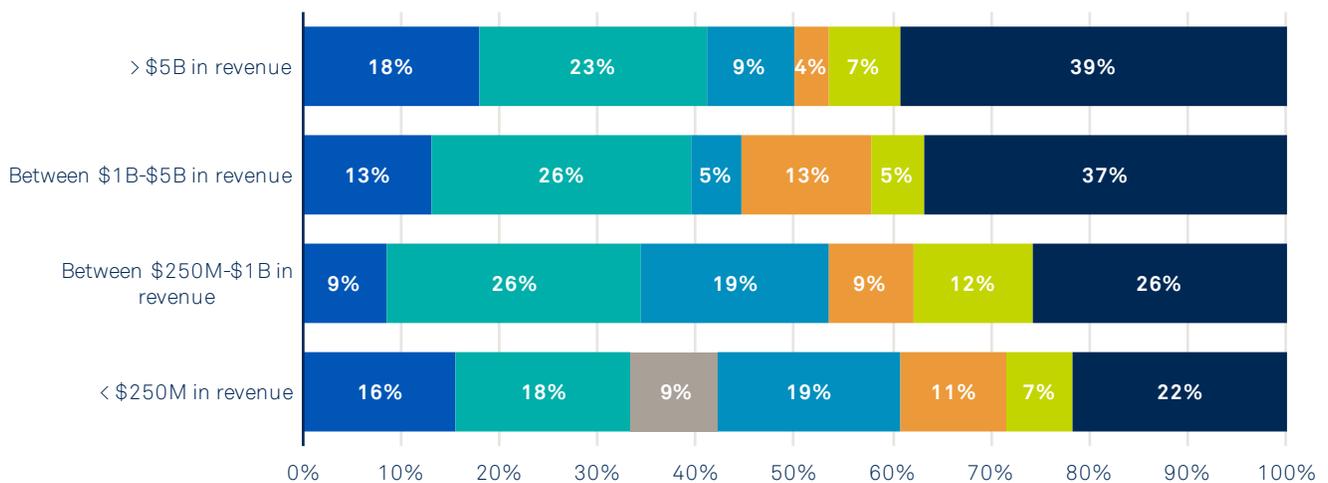
In larger companies, minimizing regulatory risk was also a key driver for 49% of respondents.

Within the smallest companies, focus is spread almost equally between reducing trial timelines (22%), reducing trial costs (19%), and minimizing regulatory risk (18%).

The need to reduce trial costs was mentioned more frequently in small and medium-sized companies (19%), compared to the largest companies, where it was selected just 5% of the time by companies with revenues between \$1 billion and \$5 billion and 9% of the time by the largest companies.

We do not believe that this suggests that larger companies care any less than smaller companies about the cost of trials. It is a universal concern. But for smaller companies—especially for pre-revenue biotech companies—trial costs are one of their largest expenses and mismanaging them can cause the company irreparable financial harm.

Figure 7: Strategic Priorities by Company Size



## What Stands in the Way?

---

The survey provided participants with the opportunity to supply free-form answers in listing the top three challenges they face in achieving their strategic priority. In synthesizing their answers, we found several common themes, mentioned in the following order of frequency:

- Improving Patient Recruitment/Enrollment
- Managing Data Quality
- Improving Patient Retention
- Containing Costs
- Achieving Timelines

This general ordering is consistent with Medidata’s experience in working with industry leaders and our understanding of the issues they face. Let’s take a look at the first three challenges in more detail.

**Within your top priority, what are the top three challenges you are trying to address (e.g., making trials more patient-centric, improving patient recruitment, improving patient retention, managing data quality, predicting adverse events, demonstrating real world efficacy, managing site relationships, developing more targeted therapies, etc.)?**



## Challenge#1:

# Improving Patient Recruitment and Enrollment

---

If the need to improve patient enrollment isn't top of mind for every clinical operations executive, it ought to be. [Seventy percent of trials are delayed by more than a month](#), and recruitment is the leading cause of those delays. In order to meet their recruitment goals, [most companies must nearly double their original timelines](#). Said another way, *patient recruitment typically takes almost twice as long as sponsors expect*.



*The integration of patient recruitment and retention into the continuum of care is very early stage, but there are numerous signs that the clinical research enterprise must go in this direction to remain viable.*

---

**Ken Getz** | Director of Sponsored Research Programs at Tufts Center for the Study of Drug Development

These delays quickly escalate costs—both in the operation of a trial and in the lost opportunity of not getting to market. Delays also may also have subtler, but far-reaching implications.

The start of a clinical trial is a very public affair, with the world—and financial analysts in particular—watching closely. Analysts may see slow enrollment as a proxy for the asset's eventual performance on the market, speculating that the issue stems from a poor comparison to the standard of care or from the treatment burden placed on patients.

As hard as it is to recruit patients today, the problem will most certainly get worse before it gets better. With so much of R&D concentrated in oncology and rare diseases, competition for a finite pool of patients will increase. And, with the rise of precision medicine, patients will have to meet inclusion/exclusion criteria to participate, that may include presenting with a very specific and possibly rare genetic profile.

The current approach to recruitment—distributing the task to investigators—puts a complex problem into the hands of people who are not prepared, nor empowered to solve it. Investigators' training and focus is on treating patients, not on recruitment advertising or promotion. The issue calls for a marketing solution that relies on an ecosystem of data, analytics, technology and services, at a minimum.

## Challenge#2: Managing Data Quality

---

The fact that “managing data quality” surfaced as the second most common challenge in the survey is logical for two reasons. First, it is one of the keys to realizing the top priority of senior executives: achieving regulatory approval. Second, the breadth of available trial data is expanding exponentially and strains existing systems and processes. [Since 2000, the number of trial sites outside of the U.S. performing studies has doubled](#), and the clinical data spectrum has expanded to include data from electronic data capture (EDC) systems, medical imaging, labs, electronic health records (EHR), genomics data, and biomedical sensors. The volume and rate of capturing these data are several orders of magnitude greater than traditional clinical data sets, and the old methods of capturing, cleaning, and analysis are simply inadequate.

The basis for high-quality data, of course, rests in having well-designed case report forms (CRFs) and well-trained site staff. But there is far more that can be done, and indeed, progressive companies have begun to use technology and analytics in creative ways to dramatically improve clinical trial data quality.



“  
*Data is the currency of  
 this industry*  
 —  
**Daniel Kraft** | Chair and Curator of  
 Exponential Medicine, Faculty Chair  
 Medicine, Singularity University

## New Data Management Paradigm

---

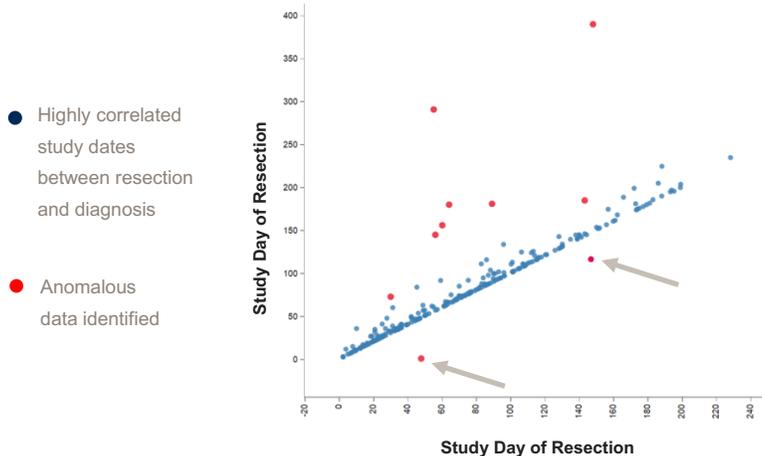
While clinical data management and monitoring have traditionally looked at the patient one-dimensionally—treating data from each source as a separate view of the patient—the patient can now be seen more holistically. This requires that we move from static data management (entering, reviewing, and locking data in succession) to adaptive data management (ADM). ADM inserts the discipline of statistically cleaning trial data by leveraging machine learning and advanced algorithms. Iterative and adaptive, such a process is system—not people—driven.

In this new data paradigm, software compares every variable in the data set to every other variable, looking for and identifying statistical relationships between them. The system can then flag outliers for investigation. And, patterns can be detected as the data accumulate, highlighting “problem” patients and/or sites that may require investigation and intervention (Figure 8).

Adoption of such tools will, we believe, begin to merge all of the functions tasked with data quality—including data management and monitoring—into a single, centralized function.

Machine learning can then also be used to guide risk-based monitoring and to identify the best sites for a given trial, based on their past performance across various data quality metrics.

Figure 8: Outlier Detection Using Machine Learning



In this Phase III oncology study, a relationship was found between when a tumor was diagnosed and when it was removed. The red dots indicate unusual instances, such as when the tumor was removed several months after diagnosis, or as indicated by the arrows, before the tumor was even diagnosed! These outliers were found even after the data had been finalized, cleansed, analyzed, and submitted through traditional means.

## Challenge#3: Improving Patient Retention

Although estimates vary, [the loss of enrolled patients in clinical trials, known as patient “leakage,” has been estimated to be roughly 18%](#). The loss of enrolled patients is, of course, a threat to the validity of trial data and a key focus of the U.S. Food & Drug Administration (FDA). The agency has drafted guidance [to help companies avoid introducing bias as a result of patient withdrawals](#), and statisticians must perform sensitivity and tipping point analyses to see if the study conclusions would be the same with and without the missing data.

Any amount of leakage is more than just unfortunate, and not just for statistical reasons. After a company has worked so hard to recruit patients and successfully screened and enrolled them, losing even a small number is a crime of inefficiency.

To compensate for anticipated leakage, companies sometimes over-recruit—in effect spending more than would be necessary if they could achieve higher retention rates. [The cost of recruiting a patient is 12 times greater than the cost of incentives and communication tactics used to retain them.](#) However, the driver in retaining patients is scientific, not financial.

## How Much Do You Rely on Data and Analytics?

Although the life sciences industry has been slower than other business sectors to adopt data and analytics to improve business decision-making, almost all respondents (90%) rely on some form of analytics to achieve their strategic priorities (Figure 9).

These responses are more revealing when broken out by company size (Figure 10).

Of the few holdouts (12 companies) that are not taking advantage of data and analytics, most (nine) are small companies with revenues <\$250M. Small companies whose business is concentrated in one or two therapeutic areas may well not have a need for external data; they may, themselves, be the leading experts in their specialty. Their use of in-house analytical solutions, however, may be driven by budget.

Meanwhile, most of the larger companies (85%) use in-house data and purchase analytical tools and services from external partners. This stands to reason, as larger companies will have amassed a large internal dataset from having run trials for years. On the other hand, analytical capabilities are not necessarily collected over time and may easily be outsourced.

Figure 9: Reliance on Data and Analytics

To what extent are data and analytics currently being used in your organization to address your top two priorities? (select one)

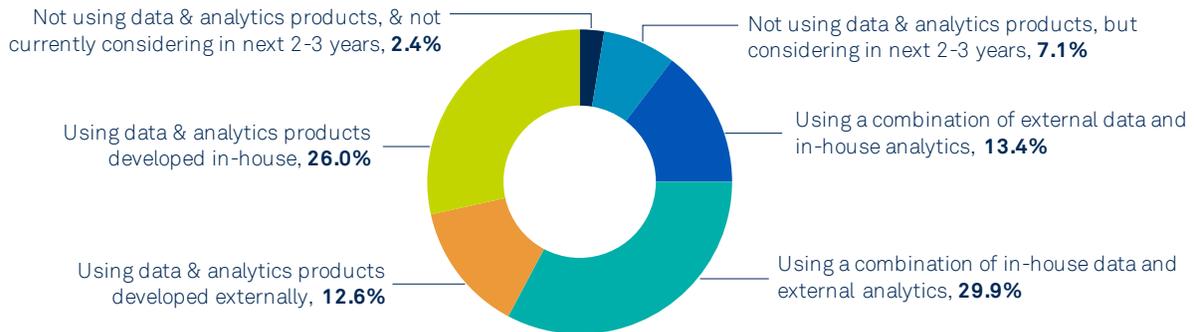
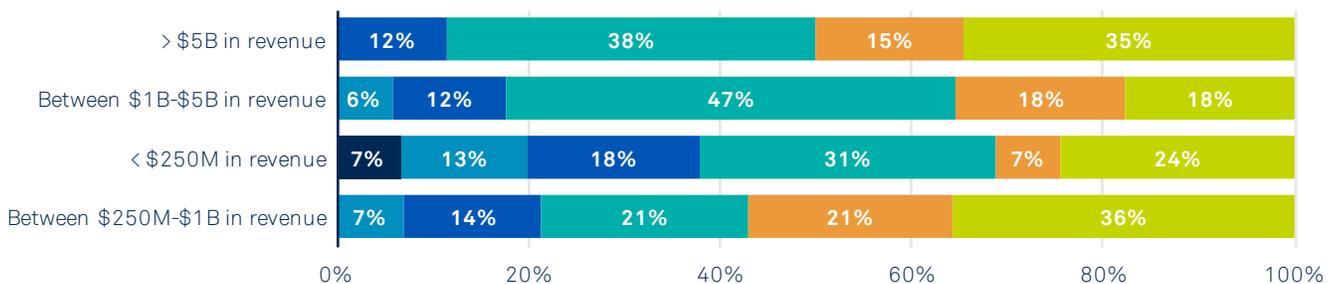


Figure 10: Reliance on Data and Analytics by Company Size



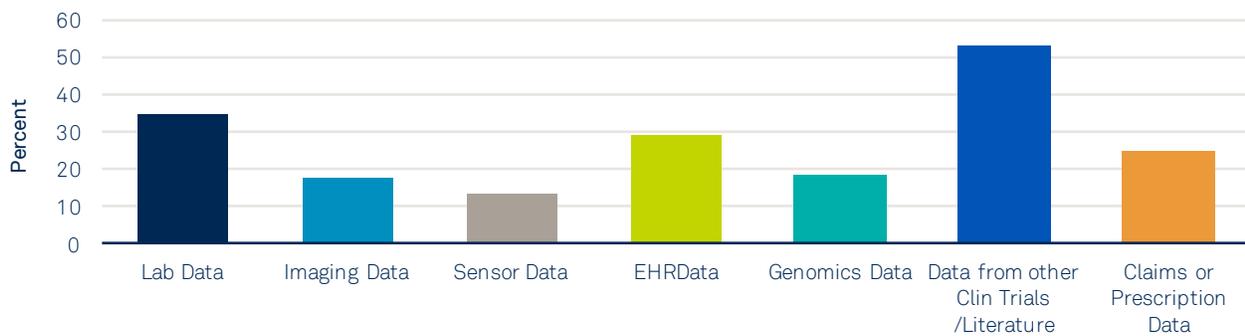
## What Data Could Help You Achieve Your Priorities?

A glance at Figure 11 reveals that respondents rely heavily on data from other clinical trials and data reported in literature. Data on historical trials is, of course, critical to understanding the science of discovery and what development approaches have been successful. Insights from historical clinical trials can inform future trial plans. Beyond that, respondents' data needs are diverse and range from traditional data types such as lab and claims/prescription data to newer sources such as EHR data, biomedical sensor, and genomics data. If we combine the mentions for EHR and medical claims/prescription data, such RWE clearly becomes the type of information that respondents most often rely upon.

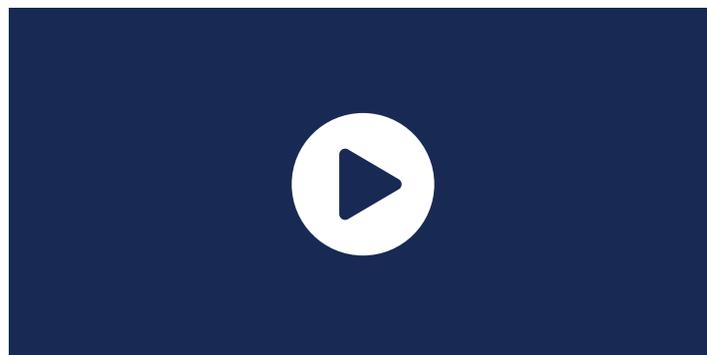
It is noteworthy that mentions of biomedical sensor data and genomic data are both relatively low. We foresee that this will change as technology develops; data from sensors and wearables will play an increasingly important role, particularly in certain therapeutic areas; it is just not there yet. And, genomic data is more important in some therapeutic areas than others, so would not be universally needed.

Figure 11: Data Needs

For the first priority you outlined in an earlier question what data could help you achieve that priority? (select all that apply)



Watch the video, The Future of Big Data.



## What Analytics Could Support Your Top Priority?

Respondents list a wide range of analytical techniques that could be useful to them in achieving their goals. Clearly, participants place a high value on predictive analytics, trial design optimization, performance benchmarking and tools for site selection—steps that can reduce trial timelines and costs.

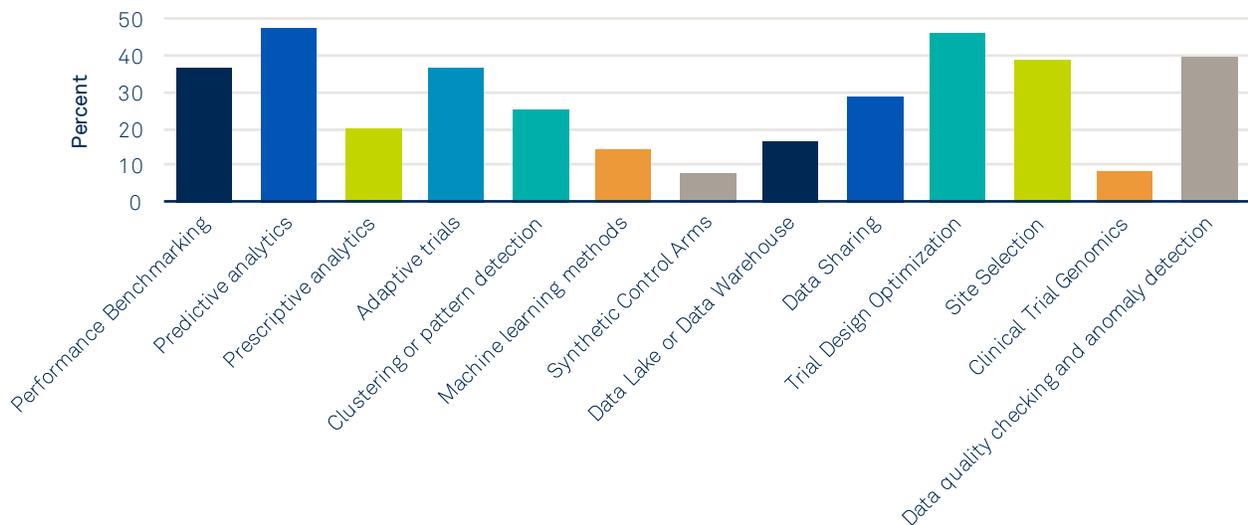
Automated techniques for improving data quality ranked highly as well, including the use of algorithms to comb through trial data for patterns and to spot anomalies.

Interestingly, use of synthetic control arms (SCAs), or historical control data, was mentioned infrequently. This perhaps was because the term is not well understood by survey respondents—a hypothesis supported by the fact that SCA is directly related to data sharing, which was mentioned much more often.

An SCA is a technique by which archived data from multiple analogous trials serves as a control arm for a current trial that, for practical or ethical reasons, cannot have its own control arm. Creating SCAs is one of the primary benefits of having access to a rich history of data from multiple sponsors across therapeutic areas.

Figure 12: Technology/Methodology Needs

What technologies or methodologies could help you achieve your first priority? (select all that apply)



Read about the use of [Synthetic Control Arms in a particular phase I/II single-arm trial in Acute Myeloid Leukemia](#).

## Conclusion

---

The survey has gathered insights from a sampling of companies across the industry and provides a glimpse into what matters to their leaders in achieving clinical development goals. Their responses confirm a growing sense that staying on time, on budget, and delivering high-quality results with clinical programs is becoming harder to do.

The volume of data in healthcare is growing at a massive rate as we wade into RWE, omics, and high-frequency sensor data. But such big data can only come to the rescue if it is properly integrated, standardized, and analyzed to provide actionable insights. And, of course, scale alone won't be enough. Data must be of high quality; the cost of bad data could be extremely high should it drive the wrong decision.

The surveyed companies are at different stages in applying the latest advanced solutions. We anticipate that those companies that can successfully leverage new data sources and new algorithmic techniques for decision-making will be in the best position for their go-to-market strategies.

We at Medidata have a vision for the future in which clinical trials maximize the use of data science to improve decision making, increase speed and quality, and reduce risk and cost.

To learn more about how Medidata is working to power smarter treatments and healthier patients, visit the [Medidata Optimize Outcomes page](#).

## About Medidata

Medidata is reinventing global drug and medical device development by creating the industry's leading cloud-based solutions for clinical research. Through our advanced applications and intelligent data analytics, Medidata helps advance the scientific goals of life sciences customers worldwide, including nearly 850 global pharmaceutical companies, biotech, diagnostic and device firms, leading academic medical centers, and contract research organizations.

The Medidata Clinical Cloud® brings a new level of quality and efficiency to clinical trials that empower our customers to make more informed decisions earlier and faster. Our unparalleled clinical trial data assets provide deep insights that pave the way for future growth. The Medidata Clinical Cloud is the primary technology solution powering clinical trials for 17 of the world's top 25 global pharmaceutical companies and is used by 16 of the top 20 medical device developers—from study design and planning through execution, management and reporting.

[info@mdsol.com](mailto:info@mdsol.com) | [mdsol.com](http://mdsol.com) | +1 866 515 6044

### Medidata Clinical Cloud®

Cloud-based clinical research solutions | Innovative technology | Data-driven analytics  
Reduced costs | Improved time to market | Faster decisions | Minimized risk