

2026 Clinical Data Trend Report



Table of Contents

Outlook for 2026: New Forces Driving Impactful Innovation	3
Trends reshaping clinical data	4
1. Risk-based data management theory needs to become reality.....	4
2. Automated orchestration will become core to the clinical data management role.....	6
3. Smart automation and meaningful AI are ready to deliver ROI.....	7
4. We're moving from 'all-encompassing' MDR to leveraging core standards better.....	8
5. Study design is incorporating patient optionality and site capability.....	9
6. Sponsors and sites are working to remove transcription.....	11
Emerging trends to watch	12
7. ICH E6(R3) demands innate traceability and drives proportionality.....	12
8. Endpoint-driven design enables RBDM.....	12
9. Deeper focus on eCOA data reliability and integrity.....	13
10. Boosting data management resilience and future economics	14

Outlook for 2026: New Forces Driving Impactful Innovation

In clinical data, there is often a lag between a new technology becoming available and its widespread adoption and return on investment (ROI). For instance, it took about a decade before most of our industry regularly used electronic data capture (EDC) during clinical trials instead of paper.

Instead of signifying caution about embracing innovation, these delays show that it takes time to develop new processes and embed change management, and to overcome risk aversion in our highly regulated industry. Today we believe there are six key trends to follow closely, each of which already has real use cases. Along with four emerging trends, we anticipate they will reshape clinical data in years to come.

A unifying theme across these ten trends is the notion of “simplify and standardize”. This applies focus to the increased complexity in the clinical landscape and drives us toward impactful innovation that gives us more value with less risk.

We are seeing greater connection between cross-functional teams, and many of these trends are themselves interconnected. For example, the shift to endpoint-driven design is foundational to the rise of risk-based data management (RBDM). Others, such as the pivot to smart automation and ramp up of AI initiatives, as well as the changing perspective on metadata repository-driven (MDR) builds, reflect where companies want to focus their energies. Many used AI pilots to test, fail, and learn and now believe that a mix of rule-driven and AI-based automation will deliver the most significant cost and efficiency improvements.

We hope this report helps your teams ideate, plan, and prioritize their clinical data initiatives so that we can deliver better trials for all.



Drew Garty

Chief Technology Officer
Veeva Clinical Data

Six trends reshaping clinical data

1. Risk-based data management theory needs to become reality

Regulators have long encouraged risk-based approaches to quality management (RBQM) and are now applying the same principles to data management and monitoring. Risk-based monitoring is advancing, but clinical data management (RBDM) is lagging behind, with less industry and regulatory maturity despite the availability of guidance.

Over half of data managers expect RBDM to impact their role in the near future **[Figure 1]**, but most are yet to make the leap of faith and move away from the security blanket of comprehensive review models. Given ever-expanding data volumes, it is not sustainable for biopharma companies to scale data management linearly using traditional methodologies. Data management leaders are now looking at ways to institutionalize RBDM as an industry.

FIGURE 1

**Data managers' response to the survey question:
"How do you see your role and its processes evolving over the next two years?"**



of data managers believe
RBDM will be involved
in their role progression
over the next two years

Source: *Clinical Data Industry Research 2025*

This will require data managers to become strategic partners to other functions, such as RBQM, and utilize smart automation and AI to eliminate manual effort and stale data (see '**Automated orchestration will become core to clinical data management**' and '**'Smart automation and meaningful AI are ready to deliver ROI'**'). These will be required if we are to move towards RBDM.

REAL USE CASES

Some sponsors are experimenting with historical trend data for proactive issue management. After defining thresholds and sharing data across departments, they assess how a trend changes over time, communicate their findings, and document issue remediation.

To do this, you need to run fit-for-purpose methods from your previous trials. Then, align as a cross-functional team on critical risks and gather input early from your study team [Figure 2]. Some of the risks you identify can be tolerated in the trial as long as you have mitigation plans and procedures in place. Once the trial starts, empower a centralized team to review and monitor data: they will be able to identify signals and data anomalies as they come in, and surface information to the right people.

“

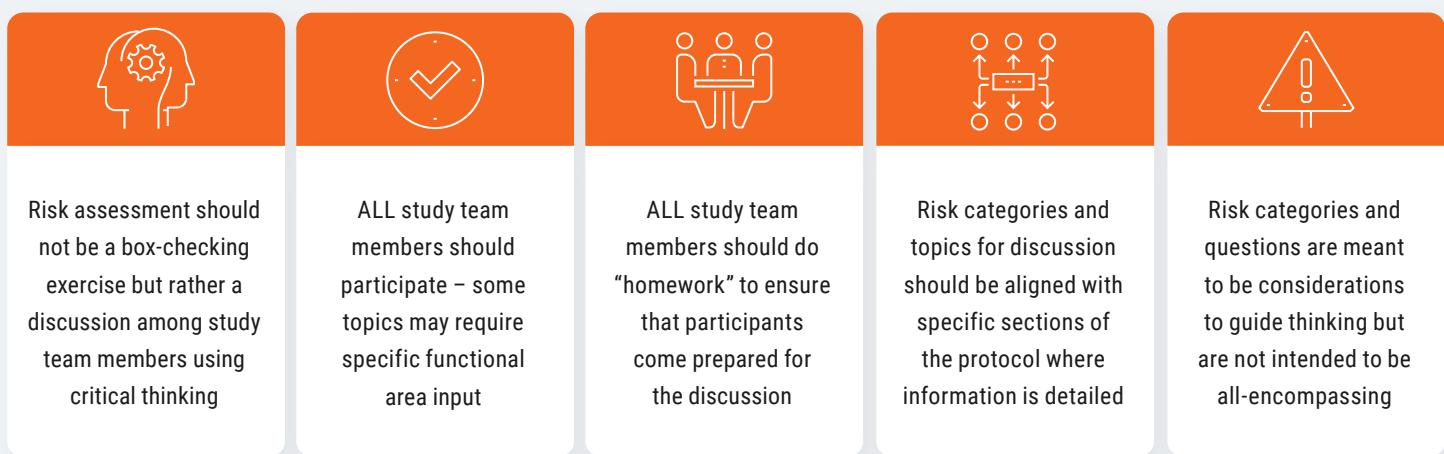
For me the biggest challenge is the resistance to change. We need a change in mindset from cleaning everything to focusing on what we need to clean. My dream would be to set up my study in EDC, feed that to CDB, and based on what we have done in the past, have AI categorize critical data and recommend listings and checks and come up with a proposal for a risk-based plan”

Valerie Balosso

Senior Director Data Strategy and Management, GSK

(Statements made by the contributor are their own opinions and do not necessarily reflect those of the company.)

FIGURE 2
Structuring a productive risk-mitigating discussion



Source: CluePoints

RBDM introduces value-creation opportunities to your trial:

1. **Higher data quality, faster** (due to proactive issue detection) leads to accelerated approvals
2. **Greater resource efficiency** (through centralized data reviews) reduces trial costs
3. **Shorter study timelines** (from reduced time to database lock) speed time to market

GSK is working on operationalizing risk-based approaches, which requires data quality categorization to pinpoint which data to focus on. The sponsor is using CluePoints central statistical monitoring tools, for example to follow up on specific sites that require attention. GSK also uses Veeva CDB to contribute to risk-based approaches, facilitated by real-time data access, data integration with eCRF and all external data, and visualization tools like the clean patient tracker.

2. Automated orchestration will become core to the clinical data management role

The data manager's role has been evolving for years. Data management has always been a collaborative endeavor, getting data to medical reviewers and other stakeholders when and where they need it, and assuming responsibility for data that has been deemed complete and accurate.

This type of orchestration is not new to data managers, but it will become more important in their role as focus shifts to RBDM and automated tools are more widespread. Automating orchestration will allow data managers to enlist AI agents to be productive on their behalf. For instance, reducing clinical programming effort by automating CQL generation from natural language.

Brian Johnson, former head of R&D technology at Takeda, comments on his experience of communicating technology and role changes: "We need to bring people along to understand that while we might be deemphasizing their role in the value chain, the value chain as a whole gets better. I might have to change what a data manager does in their day-to-day operations, but in turn, we will avoid having database locks and unlocks. It's being able to describe the value in such a way that people can understand the changes or the sacrifices that they may have to make."



[Watch the full episode](#)

EMERGING USE CASES

Fast-moving dose escalation studies require medical experts to identify signals in near real-time data. One top 20 biopharma's data management team is ensuring these medical experts have access to the data they need in a fit-for-purpose way. They also have increased responsibility in ensuring the reviews are completed and documented.

3. Smart automation and meaningful AI are ready to deliver ROI

Inefficiencies are embedded in clinical data management workflows.

According to **industry research**, too many manual steps, inefficient workflows, and disconnected systems are the main reasons that data management tasks take more time or effort than they should. Automation has been a core focus for the last few years to speed up time-consuming tasks like manual reconciliation, data cleaning & review, and query management. As automation continues to deliver ROI, companies are getting ready for the next wave: smart automation.

'Smart automation' aims to leverage the best automation approach – whether AI, rule-based, or other – to optimize efficiency for each specific use case. Importantly, it prioritizes risk management and thoughtful applications of AI, putting value creation ahead of a shiny label.

As we review the impact that AI has delivered against its promised value over the last year, skepticism from data managers has increased. Smart automation tools that lower technical barriers for data managers and help them move away from the 'air traffic controller' type tasks towards orchestration will be viewed favorably.

EMERGING USE CASES

Of the AI-augmented solutions that are more likely to succeed in delivering value in the near term, medical coding stands out as a clear leader. AI augmentation fits neatly into a slightly modified medical coding workflow **[Figure 3]**. Traditional rule-based automation is already in place today and accounts for most of the automation of medical coding. For records that do not get automatically coded, AI can be applied to either offer a medical coder a suggestion or to automatically code and have the medical coder review the selected term.

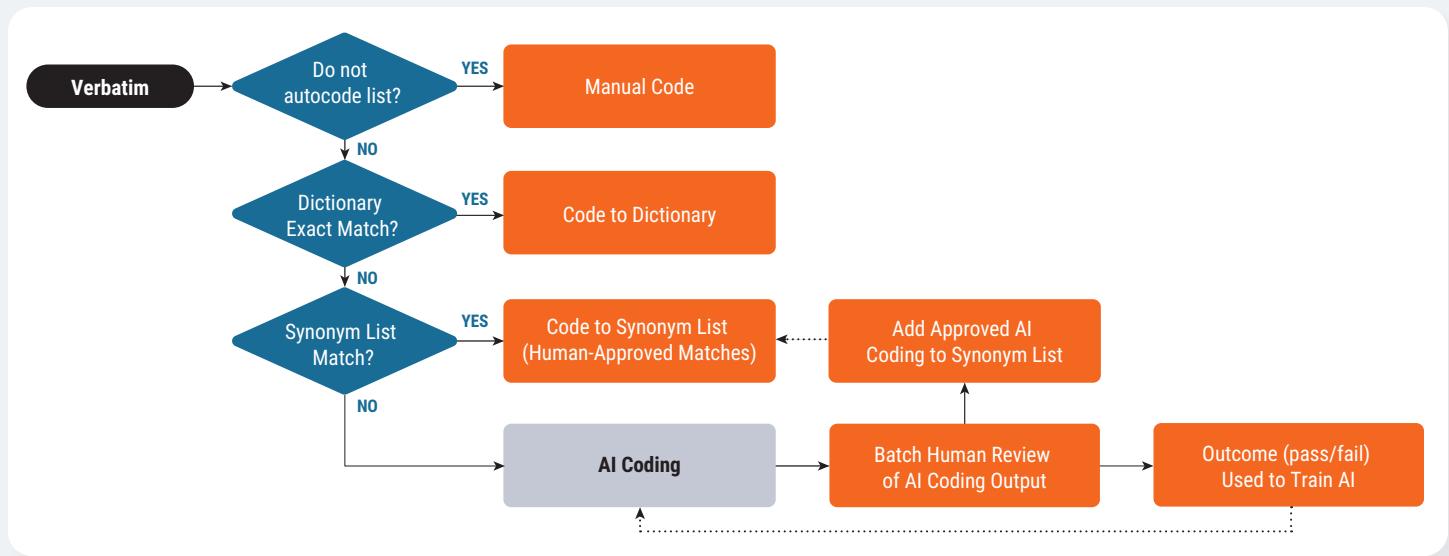
“

We're doing things with AI around testing and validation of our systems. We're currently working with Veeva on using AI to actually set up our systems. Then the next steps would be trying to handle the data that we collect in clinical trials and replacing the manual effort there with AI algorithms. I think this should very much be possible, but we need to get the experience working with AI while not influencing the impact and outcome of a clinical trial."

Ibrahim Kamstrup-Akkaoui

Vice President, Data Systems Innovation, Novo Nordisk

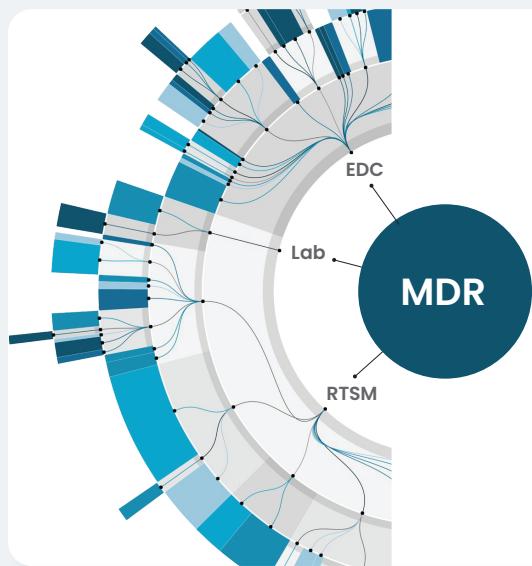
FIGURE 3
AI-augmented medical coding



Source: Veeva Systems

4. We're moving from 'all-encompassing' MDR to leveraging core standards better

FIGURE 4
Focusing MDR on what matters



- ✓ Design studies with the end in mind (analysis)
- ✓ MDR is an important solution to tie together:
 - study design
 - data collection
 - analysis
 - submission
- ✓ Study design should start with MDR and (through relationships) define data collection and analysis
- ✓ Data collection design should be as automated as possible
- ✗ MDR should be a repository of all (or as much) data collection metadata to automate collection build

Source: Veeva Systems

Metadata repository (MDR) solutions help tie together study design, data collection, analysis, and submission. Once EDC became the principal system in clinical data collection, the prevailing view was there should be one repository for all (or almost all) data collection metadata to automate study builds. In reality, it proved challenging for companies to scale metadata management, particularly because many still rely on spreadsheets.

REAL USE CASES

Instead of a repository for all metadata, a more effective emerging strategy is to focus MDR on what matters: the study design metadata that are common, shared, and critical to data management and statistics. For instance, when looking for common study design metadata between data collection and data analysis, there might be 25 properties (out of more than 1,000) of EDC metadata that affect downstream programming and analysis. The study design would start with MDR, and agree standardized data definitions at the data collection stage.

Data management and stats could then work in parallel on delivering to the same definition.

Moving from 'big', all-encompassing MDR to simplified standards will accelerate the path from study build to database lock. Traditional MDR and spreadsheets slow down companies whereas a pragmatic approach will mean they can deliver value faster. For example, Faro Health, a generative AI company, can now automate the creation of an EDC study build in just seven API calls.

EMERGING USE CASES

There is a debate surfacing on whether data standards are becoming more important, or whether technology will make them obsolete. Adopting biomedical concepts is the only way to eliminate legacy standards; biomedical concepts add both standardized data elements and the relationships between those elements and give them clinical meaning. This allows organizations to automate data analysis and clinical interpretation of data.

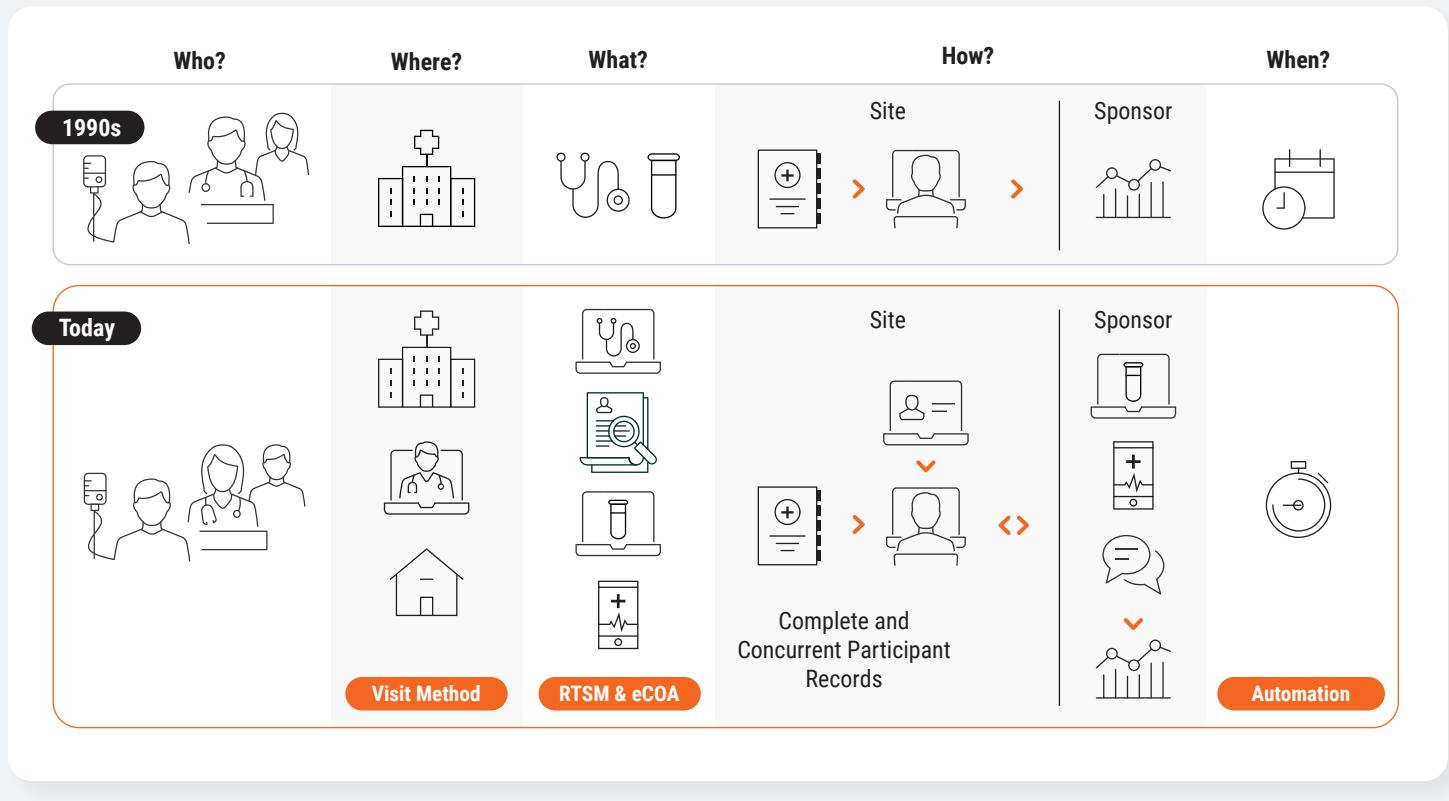
5. Study design is incorporating patient optionality and site capability

In recent years, trials have evolved from every patient visit occurring at a site to greater visit method optionality within protocols **[Figure 5]**. So there's growing demand for systems that allow sponsors and sites to provide trial options for patients more effectively. Sponsors are becoming more open to use their specific capabilities, such as decentralized methods, to improve the patient experience.

Study designers can now record whether patients were given visit method options directly in Veeva EDC using Visit Method; having this optionality built into the system will help sites reach more diverse patient populations, but delivering patient choice increases the pressure on sites and sponsors to handle more complex data. This will make access to connected technologies all the more important – from monitoring subgroup enrollment and retention, to data management and statistical analysis.

FIGURE 5

Trials are unrecognizable from 30 years ago



Source: Veeva Systems

EMERGING USE CASES

Some clinical data leaders are alleviating the patient burden at the protocol design stage, by asking study participants for less data. Others are considering whether there are tangible patient benefits before introducing new applications (e.g., eConsent) or using surveys to understand the patient experience and pinpoint improvements.



[Watch the full episode](#)

“

Sometimes you have to not only put yourself in the site's place, but the patient's place. There's a patient burden that goes beyond the site visit itself which has to be respected. So when they arrive at the site, everything's ready and what you're doing is absolutely needed. That's where things like ePRO can come in. I would love to see more protocol optimization, really mapping out every single visit.”

Joyce Moore
former Global Head of Patient Engagement, Allucent

6. Sponsors and sites are working to remove transcription

Modern trial methods require sponsors to aggregate digitized and non-digitized information from lab feeds, diagnostics, medical devices, and electronic health records (EHR). While clinical workbenches centralize data cleaning, aggregation, and reconciliation for sponsors, sites currently lack access to a centralized database for all their relevant clinical data sources.

For example, EHR integrations may be a viable trend for the largest sites but have been challenging to scale and adopt. In 2026, organizations will take meaningful steps to solve EHR/source integration at scale.

Instead of mandating technology for sites to use, regardless of their size or resources, some sponsors are now giving sites the choice.



[Watch the full episode](#)

“

What I've seen is a shift in the industry regarding site centricity. Sponsors actually mean it now, it's not just a talking point. I see the actual shift, sponsors letting sites pick the technology that works for them. Everybody benefits if you support their use of it.”

Joe Lengfellner

Senior Director, Clinical Informatics & Technology,
Memorial Sloan Kettering Cancer Center (MSKCC)

REAL USE CASES

Quicker queries are one of many innovations that will simplify all sites' experience by reducing the time and effort spent on a time-consuming, manual activity: reviewing and responding to queries. CRAs will no longer have to type common messages and responses: if they see a bad value, they can query that value in one click (similarly, sites would verify, amend, or respond to queried data points in a single click).

One of the ways Alcon, a medical device company, evaluates the site experience is to assess how quickly sites complete data entry: a longer lag could indicate systems and databases are not user-friendly. A platform-based approach is helping Alcon support site needs. Leianne Ebert, head of clinical data operations, explains: “This is something we monitor regularly, and last week, our records showed that 45% of our data is entered on the same day as the visit date.”

Emerging trends to watch

7. ICH E6(R3) demands innate traceability and drives proportionality

The new ICH E6(R3) guidance encourages more flexibility in both protocol design and study activities to simplify the patient experience, but that flexibility comes at a cost. Sponsors must demonstrate that they are going to carry out all the necessary actions for oversight that they claim they will do, down to a data point level. For this, they need traceability built into the system.

ICH E6(R3) also drives the notion of proportionality of activity. We should be focusing resources on the most important tasks to maximize their value potential – a more measurable dimension than efficiency gains alone. Proportional oversight will become an important outcome of the new guidance, with exploratory data requiring more relaxed oversight and freeing up sponsors to focus on critical data and processes.

8. Endpoint-driven design enables RBDM

How can risk-based data management (RBDM) theory be applied in practice?

Clinical research is fundamentally a statistical endeavor, but statisticians are time-poor and typically do most of their work at the backend after data collection. But there are some indications that teams such as statistics and data management invest too much time in activities that do not proportionately improve the quality of the final database used for analysis: for instance, reviewing, verifying, and querying data that rarely changes.

Endpoint-driven design enables RBDM by proactively defining critical data and focusing data cleaning effort on those endpoints.

If statisticians were involved sooner in data management, they could help define primary and secondary endpoints independently of data management [Figure 6]. They would map out the required data, rank it, and identify what's critical (and what isn't). With clearer data cleaning expectations, data management will know which data to focus on and can create queries that advance the study.

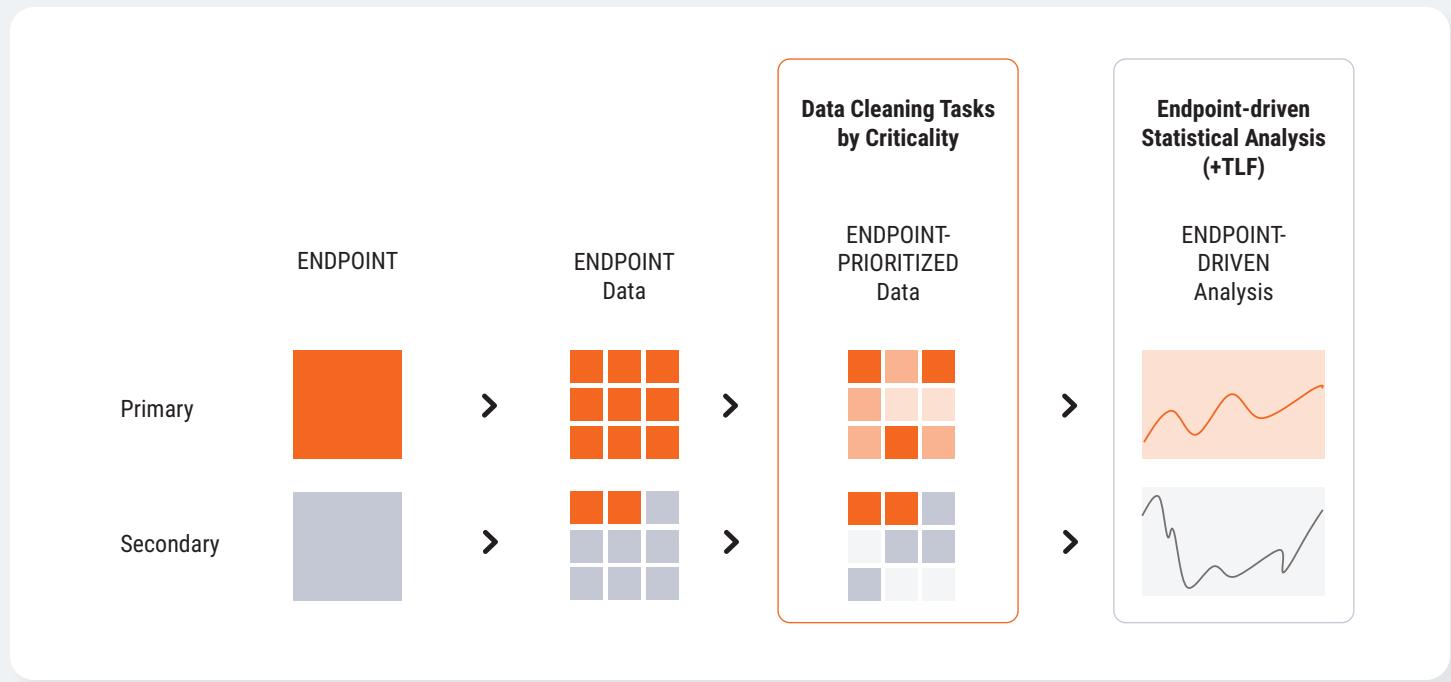
Involving stats earlier would improve data management. These teams will question whether endpoint data is missing or unlikely to be used. Over time, endpoint-driven design will lead to less single-use and unnecessary data being collected.

“

We can end up creating more noise and problems for ourselves downstream if we haven't thoughtfully introduced quality-by-design upfront and risk-based approaches to how we collect, curate, and clean data. If we can make the argument about the criticality of data collection that's contributing to endpoint-driven study design, then I think we can have a louder voice at the table.”

Head of Data Management,
Global Biopharma Company

FIGURE 6
Bringing stats to the forefront of data management



Source: Veeva Systems

9. Deeper focus on eCOA data reliability and integrity

Innovation has been front of mind for eCOA and continues to be a priority in the industry. But it is imperative that the data meets regulatory standards. Despite years of digital transformation, many sponsors still spend disproportionate effort checking and cleaning data because they can't trust it. Unreliable data leads to increased time, effort, and risk to study results. Regulators and sponsors alike are demanding stronger traceability and transparency across the data lifecycle.

A growing dissatisfaction with the status quo is turning attention towards modern eCOAs as part of a unified clinical data platform. These systems standardize eCOAs across studies using a validated framework to reduce

variability and simplify build. On-demand real-time data enables early issue identification, leading to proactive and confident decision-making.

The organizations leading the next phase of innovation are those removing risk through system design from the outset, not reacting to it downstream. This will rely on designing eCOAs to maximize data completeness as well as reliability. For example, designing the system to ensure all data points are entered and identifying missed entries before they become a repeat problem. This streamlines both data management and clinical operations by removing endless loops of checking in the field by clinical research associates (CRAs).

10. Boosting data management resilience and future economics

In recent years, clinical research has continued despite wars, pandemics, and unrelenting macroeconomic pressures. While the context varies, the scrutiny of value delivered, patient outcomes, and resource efficiency remain consistent.

There is currently some uncertainty across several dimensions in the industry: from the direction and leadership of the FDA to the M&A approach favored by the Federal Trade Commission (FTC), to drug pricing

negotiations and more. In addition, we may see a need for more post-market efficacy studies that bring new data challenges.

Data management will need to become more resilient to address these risks proactively. As the discipline transitions towards orchestration, it can play more of a strategic role by identifying critical data points, getting involved earlier in the conversation, and educating key decision-makers on the results.

As the clinical data landscape encourages us to simplify and standardize, embracing these ten trends will be essential for biopharma companies to deliver faster, higher-quality trials.

Learn more about connecting your data flow.



ABOUT US

Veeva is the global leader in cloud software for the life sciences industry. Committed to innovation, product excellence, and customer success, Veeva serves more than 1,100 customers, ranging from the world's largest biopharma companies to emerging biotechs. As a Public Benefit Corporation, Veeva is committed to balancing the interests of all stakeholders, including customers, employees, shareholders, and the industries it serves.

For more information, visit [veeva.com](https://www.veeva.com)