Paper DV10

Detecting Trends Before They Appear: AI-Driven Visualizations in Clinical Trials

Steven Tan, Wu Consulting Group, Boston, United States
Harrison Gu, Wu Consulting Group, Baltimore, United States

ABSTRACT

Early detection of trends and anomalies in clinical and safety data is critical for timely decision-making. As such, the implementation of AI-driven data visualization tools is an emerging area of promise. This paper investigates the strategy shift required in the clinical trial industry following the post-COVID technology adoption plateau, which sees a decreasingly rate of technological implementation. To address this, the paper presents a multifaceted framework comprised of three core arguments. First, technology adoption is highly case-dependent, which requires a carefully customized selection of the appropriate tools for the appropriate issue. A 40/40/20 portfolio approach is proposed to maximize efficiency of technology resource allocation. Second, the research justifies technology adoption for risk mitigation and competitive necessity arguments, even when the precise Return on Investment (ROI) is difficult to quantify. Lastly, the economic value of minimizing "decision lag" in the clinical trial space through the deployment of AI-driven data visualization provides sufficient reason for technology adoption. Ultimately, the paper confirms not only the necessity of technology adoption but also the strategic framework to determine the optimal technology implementation approach in the post-pandemic digital age.

INTRODUCTION

In the modern pharmaceutical landscape, the initial surge of digital adoption has reached a significant plateau, creating a critical need for a more strategic approach to AI-driven data visualization. Despite nearly 81% of clinical trial sites utilizing digital tools by 2022, actual trial outcomes have not seen a corresponding uptick; currently, 80% of trials fail to meet initial enrollment goals, and 85% fail to retain sufficient patients. This disconnect suggests that organizations may be over-investing in basic visibility tools, such as simple dashboards, while under-investing in critical problem-solving AI that addresses fundamental bottlenecks like patient-finding and protocol optimization. Consequently, the justification for advanced technology must shift from a narrow focus on immediate, quantifiable Return on Investment (ROI) to a multifaceted framework that prioritizes risk mitigation, regulatory compliance, and competitive necessity.

CHOOSING THE APPROPIATE TECHNOLOGY STACK

The Current Reality

Clinical trials have an inherent structural tension on its operational efficiency. On the one hand, scientific and regulatory rigor requires comprehensive, high-fidelity clinical data, with everything ranging from patient demographics to lab results. On the opposing side, the reality of trial sites and patient enrollment limits what is practically collectible with hard constraints like eligibility and participation barriers on the patient side and logistical limitations on the site side. This creates a fundamental conflict between the goal and the reality of clinical trials.

Despite a recent surge in technology adoption, actual outcomes have not seen an uptick. A 2023 report discovered that the percentage of clinical trial sites using digital tools jumped from 43% to 81% between 2020 and 2022 during the COVID-19 pandemic. During this period, necessity-driven digital technologies, such as remote work, virtual trials, and data sharing, were critical to maintain operations despite widespread logistical disruptions worldwide. However, contrary to intuition, the current phase of technology implementation is characterized by disappointing trial results: 80% of trials failed to hit initial enrollment goals; 85% failed to retain sufficient patients; and all clinical trials suffered from a 30% attrition rate in sustained patient participation.[1] Given this information, the question now is about asking the "why" behind the failure of recent technology adoption. It is not because people are using technology wrong, or that trial protocols are poorly designed. It is ultimately because technology cannot address the fundamental tension in the clinical trials space.

The clinical trial industry continues to purchase novel technology in hopes of finding a miraculous cure to this issue. But that is infeasible: Technology can only optimize around the constraint, not remove it. And when the industry

finally understands that is the case, they stop adopting new technology. This is precisely why technology plateaued in the post-pandemic age. While digital tool adoption at trial sites jumped 24% from February 2021 to February 2022, adoption rate remained relatively constant with a 1% change from February 2022 to July 2022.[2] The ceiling is hit.

To put simply, the current digital adoption approach lacks balance. Currently, trial organizations are over-investing in visibility tools (e.g. dashboards) while under-investing in critical problem-solving tools (e.g. patient-finding AI and protocol optimization). The real bottleneck is with patient-finding, demonstrated by the fact that almost 50% of sites enrolled one or no patients.[3] The success of prioritizing unclogging this critical gridlock can be seen from an almost 50% improvement in patient-recruitment time at Amgen, Bayer, and Novartis after they invested billions in AI scanning and patient-finding technology.[4] However, despite these lone-wolf efforts, the industry trend overwhelmingly remains on its emphasis over dashboard adoption with a nearly 80% adoption rate of dashboard systems like EDC and 45% on platforms like ePRO.[5]

It is not difficult to see why dashboards are being heavily prioritized. Through streamlining data collection procedures, EDC and ePRO can reduce the typical trial timeline by 20%.[6] At the same time, these digital dashboard systems can also improve data quality with their built-in error checks. The data quality improvement is on average at least 50%, thereby helping on study personnel to focus their attention on higher-level decision-making rather than tedious error checking.

Meanwhile, patient-finding AI adoption is critically stagnant: in 2024, of the 29% of CROs that had implemented AI/ML, 52% had been using Ai to identify patients for trials, which means that the overall patient-find AI adoption rate is roughly a mere 15.08%.[7] This highlights the mismatch between what is important and the industry reality.

A Context-Dependent Approach to Technology Stack Selection

Technology adoption in the clinical space is a highly context-sensitive process. Different companies' needs vary drastically depending on size, scope, and resources, which highlights the necessity of selecting the appropriate technology stack accordingly. Small biotech companies, with fewer operating sites (i.e., 5-15), are encumbered by logistical constraints, such as limited staffing and tight budgets, and scalability limitations. Consequently, these firms often opt for cost-effective technological solutions to maximize operational gains and steer away from high-stakes, complex technologies. For example, a 2018 study found that while three quarters of contract research organizations (CRO) had adopted risk-based monitoring (RBM), an analytics-based clinical data monitoring strategy; of the respondents that had not adopted RMB, the vast majority were "small/niche CROs and small biotech companies." (Applied clinical trials). On the other hand, large pharmaceutical organizations, which operate with significantly more sites (i.e., more than 100), prioritize robust, scalable platforms capable of processing large volumes of data across various geographical sites. Therefore, these larger companies typically have a higher desirability for advanced AI-powered solutions to optimize data visualization for operational efficacy and decision-making.

As the industry hit the post-pandemic plateau, companies now prioritize evaluating the Return on Investment (ROI), a measure of an investment's profitability by comparing gains against costs, for technology adoption. A 2025 Grant Thornton survey reports that 67% of interviewed executives ranked technology resource optimization as a key annual agenda and 63% said that a top ROI goal is to reduce operational costs, highlighting this key shift of emphasis to ROI in technology adoption.[8] This new framework focuses on adopting digital tools that improve operational efficiency, generate tangible value, and address specific bottlenecks.

This trend marks a paradigm shift from a necessity-driven approach to an ROI-driven one. In this regard, the specific AI adoption strategy varies between small biotech firms and large pharmaceutical corporations. For small biotech, digital tools are about cost mitigation: for example, Recursion, an AI drug discovery company, uses real clinical trial data and an AI pipeline to simulate trials to determine the best patient stratification. The core of this approach is to pre-emptively reduce trial costs through spearheading their efforts in finding the cost-efficient, optimal strategy before running trials.[9] For big pharma, however, their technology ROI strategy is mainly focused on improving existing operational efficiency and streamlining protocols: an example would be Pfizer's "agentic workflow," which is used to "streamline redundant feasibility surveys and automate due diligence across global sites."[10] Not only is the clinical trials industry increasingly and strategically adopting ROI-focused technology, but the precise digital strategy also differs between small biotech and large pharmaceutical companies.

PORTFOLIO MODEL FOR TECHNOLOGY RESOURCE ALLOCATION

The primary question is no longer the necessity of technological adoption, but rather the precise strategy of choosing the appropriate technology stack for the appropriate problem. Clinical trials organizations should adopt a 40/40/20 portfolio approach with to optimize resource allocation across three technology tiers. Under this novel model, the first tier, which represents 40% of the technology budget, is dedicated to mandatory compliance tools and infrastructure, including Electronic Data Capture (EDC) licenses and CDISC compliance. Another 40% portion should allocate towards the second tier, which focuses on high-ROI technologies to overcome critical bottlenecks, such as patient-finding AI pilots, eConsent platforms, and site selection analytics. The final 20% is reserved for experimentation, such as establishing a pilot program for emerging AI dashboards and advanced analytics. The experimental projects should undergo rigorous performance measurements to determine whether they should be expanded or cut.

In practice, for example, a small biotech firm managing two to three trials per year with an annual technology budget of 300,000 to 400,000 USD would be recommended to invest 120,000 to 160,000 USD to tier-1 compliance tools under this model. A similar amount should be directed towards Tier-2 high-ROI technology, with a rough breakdown of 80,000 USD on patient-finding AI pilot; 40,000 USD on eConsent platform; and 40,000 USD on site selection analytics. Lastly, 60,000 to 80,000 USD should go towards experimental projects. The model remains robust even when scaled. The allocation percentages should remain roughly similar even for big biotech organizations running over 20 annual trials with a 3 to 5 million USD technology budget per year.

COMPRESSING DECISION-MAKING: FROM WEEKS TO MINUTES

THE PROBLEM: Decision Lag is Expensive

"Decision lag" refers to the latency in the timeframe of a clinical trial question from being posed to getting answered. On average, even at top research organizations, the delay between the last patient out (LPO) to database lock (DBL) is four weeks.[11] While this timeframe may seem negligible, it is causing an increasingly significant critical lag, as almost 80% of trials fail to meet timeline targets.[12]

Economically, this decision-making latency represents a significant drain on capital in the high-stakes clinical research industry. In a fast-spaced environment, significant costs are associated with patient management, infrastructure, and the opportunity costs of industry sponsorship in clinical trials and can compound depending on trial size.[13] In fact, a single day of clinical trial delay burns approximately 40,000 USD in direct clinical trial costs and 500,000 USD in lost prescription drug revenue.[14] This implies that just ten days of delay per year could result in a 5.4 million USD waste, highlighting an essential need to improve operational efficiency and reduce decision lag. Statistical data back up this argument: in a meta-analysis of 254 empirical articles that target recommendations for reducing trial costs, 33% of them report the importance of improving operational efficiency. A solution to reduce decision lag is urgently required.

THE SOLUTION: Dashboards Compress Time

AI-driven data visualizations can provide a necessary solution to improve operational efficiency and reduce decision lag. It helps scan large datasets, identify key patterns and correlations, and create visual representations with minimal manual labor, producing clear, actionable insights. One key benefit is in enrollment monitoring, shortening monthly site performance reviews to daily and saving approximately six to eight weeks. In a survey, after applying a real-time patient enrollment dashboard for a multisite clinical trial, 77% of the 23 clinical staff respondents "felt that it was easier or much easier to use the dashboard for communication than to use direct communication."[15] Empirical evidence supports the argument that dashboards compress decision lag. In a timeline study, it was found that adopting AI data visualizations was correlated to a 3-week reduction in a 26-week trial, which represents roughly a 11.5% shortening, after Sponsor X integrates a novel Kaplan-Meier model, which produces participant progression forecasts based on real-time data and analytics (pm360).

Another advantage lies within safety signals: historically, safety signals analysis typically relied on a two-week cycle, which can delay critical insight. By enabling real-time monitoring, dashboards can identify anomalies and red lights as they surface, effectively saving two to three weeks per week. This will not only accelerate the overall trial timeline but also mitigate safety risks in clinical trials.

The financial ROI for dashboard deployment would be high. By significantly decreasing risks and increasing trial timelines, digital tools bring significant financial outcomes that outweigh initial investments. An industry report of over 450,000 trials discovers quantifies that digital endpoint adoption in clinical trials has an ROI of "4 to 6 times the cost of implementation."[16] The same report finds that specifically for phase 2 trials, the increase in eNPV ranges between 2.1 million and 3.3 million USD with an 30%-to-48% ROI per indication. In other words, for every dollar that is invested into digital dashboards, there is a 4-to-6-dollar return.

THE REALITY: Implementation is Hard

With that said, data visualizations implementation remains difficult. A significant 60% of the implementation effort is dedicated to data integration, as systems like EDC, CTMS, and labs do not operate with seamless communication. Instead, they operate using disparate data standards and proprietary APIs. When cross-platform and cross-site compatibility is low, implementation and integration is excruciatingly hard and slow. Recent industry data shows that nearly 98% of CRO and biotech respondents replied that they have major challenges with technology implementation in their clinical applications.[17]

Companies also face a decision between R and SAS as their programming language. For a long time, SAS is often perceived as the gold standard for the clinical trials industry and is commonly used for regulatory entities like the FDA. The key advantage of SAS lies within its efficiency with large-scale data management, such as working with Study Data Tabulation Model (SDTM) and Analysis Data Model (ADaM) to convert raw datasets into standardized formats for compliance and analysis in clinical trials. Being a closed source language, SAS guarantees excellent data security, as data cannot be easily extracted and manipulated. However, cost remains a drawback, as a license is required for usage, and significantly limits the language to a closed environment. At the same time, it is less flexible in terms of text mining, which is the process of extracting information from text, and graphical representations, and is generally more difficult to use, especially for non-programmer users.[18]

Meanwhile, R presents a comparable alternative. As an open-source language, R is user-friendly and highly available for free. Its wide user base gives the user access to a host of readily available packages and add-ons on platforms like GitHub and the Comprehensive R Archive Network (CRAN). It also has high flexibility and compatibility with integrated tools, like R Shiny which can be used for generating web-based graphics interfaces.[19] Nevertheless, R struggles with performing high-volume computing tasks, such as large datasets and calculations. Being open source also means that it is inherently prone to data security threats and has limited professional support.[20]

Not only are the two vastly different in the technicals but also vary significantly on the costs. While R is technically free, it frequently requires extra, costly validation protocols and IT personnel size to meet regulatory requirements.[21] For a hypothetical 30-member team, R costs 150,000 USD on regulatory compliance; approximately 30,000 to 50,000 USD on regulatory IDE; 20,000 to 40,000 USD on package management; and 360,000 USD on IT staffing. This comes to an approximate range of 515,000 to 605,000 USD annual spending. Meanwhile, SAS's costs are broken down into an upfront license fee of 175,000 USD; 150,000 USD on regulatory compliance; none for regulatory IDE and package management; and 270,000 USD on IT staffing. Per year, SAS is associated with 550,000 USD.[22] From these data, it is apparent that the specific cost breakdown between the two languages are starkly different. Therefore, the insight is that companies should adopt a hybrid approach to both programming languages, as opposed to a religious preference for either one over the other.

THE GAP: Real Data Is Needed

Regardless, the precise effect of AI-driven data visualization in the clinical trials industry still represents a major research gap. The current literature relies solely on a single published 3-week timeline study. This is especially problematic since the study lacks crucial cost comparisons and safety outcome data between an experimental group and a control group for AI data visualization adoption. Another key weakness is the lack of solid, substantive proof behind the ROI calculations, which are only rough estimates. With the lack of these key data, the industry is essentially flying blind on million-dollar decisions. To provide a more accurate conclusion, more research studies in this field are needed.

TECHNOLOGY REMAINS USEFUL EVEN WITHOUT ROI PROOF

Apart from ROI, technology adoption has a few other net benefits. First, AI data visualizations can serve as a risk mitigation mechanism. Traditional monitoring relies heavily on source data verification (SDV), which requires comparing electronic data with original source documents to ensure accuracy and completeness. This is a highly error-prone process that can result in critical trial failure.[23] To address this, AI-driven data visualization can serve as an insurance against catastrophe by utilizing machine learning to identify key outliers early on.

Another benefit is fulfilling regulatory expectations. Global regulatory bodies, including the FDA and EMA, have shifted their expectations from suggesting to requiring robust Risk-Based Quality Management (RBQM). By visualizing data lineage in real-time, sponsors can satisfy the core ALCOA+ principles—ensuring data is Attributable, Legible, Contemporaneous, Original, and Accurate—thereby making the entire trial process "inspectable by design" and reducing the friction of regulatory filings.

In an era defined by clinical workforce shortages, AI visualization also acts as a competitive necessity. These tools allow a single Clinical Research Associate (CRA) to oversee a significantly higher volume of sites without increasing the margin of error, as the system highlights only the areas requiring human attention. Furthermore, standardizing visualization frameworks across a company's entire portfolio allows sponsors to identify systemic issues and site performance trends that would remain invisible within siloed spreadsheets. AI-driven heat maps also enable real-time reallocation of resources based on patient enrollment velocity, which significantly reduces the high "rescue" costs that typically plague the final stages of recruitment.

Laslty, AI technology can have benefits that are unquantifiable. For example, while traditional balance sheets may struggle to capture qualitative shifts, the impact of high-fidelity visualization on stakeholder confidence is profound. Real-time, transparent data visuals increase the trust of investors and internal sponsors during interim analyses, fostering a more agile decision-making environment. From a scientific perspective, the ability to visualize multi-dimensional data, such as the relationship between biomarkers and patient symptoms, often leads to serendipitous insights into sub-population efficacy. This depth of insight can potentially save a drug candidate that might have otherwise failed a broad indication by allowing researchers to narrow their focus to the most responsive patient cohorts.

CONCLUSION

Ultimately, the adoption of AI-driven data visualization is a strategic imperative designed to minimize "decision lag" and maximize the efficiency of resource allocation in an increasingly complex clinical environment. While a precise ROI can be difficult to quantify, the investment is justified by its role as essential insurance against catastrophic trial failures and as a "table stake" for meeting modern regulatory and competitive expectations. Success in this transition requires a disciplined implementation strategy that matches specific tools to specific problems, measures process-oriented metrics and accepts the inherent uncertainty of benefits that do not always appear in traditional balance sheets. By moving beyond simple data visibility toward a structured portfolio-based strategy, organizations can ensure that their technology investments serve as a genuine force multiplier for clinical success in the post-pandemic digital age.

REFERENCES

1. https://www.appliedclinicaltrialsonline.com/view/clinical-trials-site-networks-success
2. *https://dt-consulting.com/dts-clinical-trial-digital-tracker-q3-2022/*
3. https://www.appliedclinicaltrialsonline.com/view/clinical-trials-site-networks-success
4. https://www.reuters.com/technology/big-pharma-bets-ai-speed-up-clinical-trials-2023-09-22/
5. https://www.openclinica.com/blog/technology-adoption-in-clinical-trials-trends-plateaus-in-2023/
6. https://www.octalsoft.com/how-electronic-data-capture-edc-can-improve-outcomes-in-clinical-trials/#:~:text=Time%20is%20one%20key%20factor,saving%20treatments%20to%20patients%20sooner
7. https://www.appliedclinicaltrialsonline.com/view/new-insights-on-the-impact-of-ai-enabled-solutions#:~:text=For%20planning%20and%20design%20activities,support%20clinical%20study%20report%20writing
8. https://www.grantthornton.com/insights/press-releases/2025/august/grant-thornton-digital-transformation-survey-pt2
9. https://www.genengnews.com/topics/artificial-intelligence/as-pipeline-advances-recursion-expands-ai-focus-to-clinical-trials/

10. https://www.clinicaltrialvanguard.com/conference-coverage/pfizers-ai-powered-feasibility-and-cost-efficiency-the-future-of-clinical-trial-recruitment/#:~:text=A%20focal%20point%20of%20the,advancing%20toward%20more%20complex%20integrations
11. https://medrio.com/blog/clinical-trial-database-locks/?utm_source=chatgpt.com
12. https://www.fiercebiotech.com/sponsored/avoiding-costly-clinical-trial-delays-takes-more-breaking-down-silos#:~:text=Clinical%20trial%20delays%20are%20a%20common%20issue,Site%20protocol%20being%20subject%20to%20major%20changes
13. https://pmc.ncbi.nlm.nih.gov/articles/PMC4746719/
14. https://pubmed.ncbi.nlm.nih.gov/30628466/
15. https://www.appliedclinicaltrialsonline.com/view/how-much-does-a-day-of-delay-in-a-clinical-trial-really-cost-#:~:text=estimate,Clinical%20trials%20in%20respiratory%2C%20immunology
16. https://dimesociety.org/newsroom/blog/deep-dive/quantifying-the-roi-of-using-digital-endpoints/#:~:text=*%20Enhanced%20efficiency:%20Digital%20endpoints%20significantly%20shorten,to%206%20times%20the%20cost%20of%20implementation
17. https://www.veeva.com/wp-content/uploads/2020/09/Veeva-2020-Unified-Clinical-Operations-Survey-Report.pdf
18. https://data-flair.training/blogs/disadvantages-and-advantages-of-sas/
19. https://myadlm.org/cln/articles/2020/april/why-clinical-laboratorians-should-embrace-the-r-programming-language
20. https://olibr.com/blog/what-are-the-disadvantages-and-advantages-of-using-r/
21. https://pharmaverse.github.io/blog/posts/2024-04-15_de-_mystifying__.../de-_mystifying__r__programming_in__clinical__trials.html
22. https://medium.com/@agunuganti_43360/the-future-of-clinical-trial-programming-balancing-sas-stability-with-r-innovation-8b4e372d8a1c
23. https://www.gcp-service.com/what-is-source-data-verification-in-clinical-trials/

CONTACT INFORMATION

Author Name: Steven Tan
Company: Wu Consulting Group
Work Phone: 630.491.4181
Email: steven@wuconsulting.com
Website: wuconsulting.com