AI Won't Fix a Broken Protocol. But It Will Expose One Faster.

What happens when we implement intelligent tools on top of a foundation that was not built to support them?

The answer is not that AI fails, but that the protocol is more likely to fail faster and at greater scale.

Protocol Quality Is Now an AI Readiness Issue

Most conversations about AI readiness focus on data infrastructure, system integration, and governance frameworks. These are necessary discussions. However, there is a more fundamental readiness question that does not get asked often enough: Is the protocol itself operationally sound enough to sit underneath an intelligent layer?

If the answer is no, you do not have an AI problem. You have a protocol problem that AI will accelerate.

In this instance, we should not confuse digitization with readiness. A protocol may look complete on paper and still be fragile in execution. Operational fragility is where protocols break down, site level and study level deviations occur, and this can ultimately jeopardize critical endpoints. Inclusion and exclusion criteria that read clearly on paper but are left to clinical interpretation create site-level variability that no AI system corrects for on its own. Human operators are accustomed to navigating this kind of ambiguity. We ask questions, seek

clarification, and make judgment calls based on experience. AI systems operates fundamentally differently. Without context and clarity confusion ensures in ambiguity. AI requires explicit logic, structured data, and clear boundaries. When an AI tool encounters vague eligibility criteria during automated patient screening or data review, it cannot interpret intent and does not have context. It stalls, miscategorizes, or generates a cascade of queries that lands back on the operational team anyway. The problem was not created with AI. The protocol was weak from the start and AI exposed it faster.

A schedule of events that was not rigorously cross-checked against central laboratory setup creates downstream data integrity issues that compound the moment automation speeds up any part of the process. AI can help detect inconsistencies across documents, and that is a meaningful advantage. However, detection is not the same as correction. If study documents were not pressure-tested before activation, AI may expose the misalignment faster and across more touchpoints, not resolve it.

Country and site mix is another area where weak protocol foundations become highly visible. A protocol may be scientifically sound and still be poorly aligned to standard of care in key countries. This creates regulatory friction, enrollment challenges, and operational complexity that no downstream automation will solve. If AI is layered in to optimize performance at that point, it will identify the bottlenecks quickly, but the technology is not fixing the root issue. It is revealing that the study was built on assumptions that were not operationally durable.

What AI Actually Does Well Here

I am not skeptical of AI in clinical trial operations. The value of AI in this context is specific, and it is worth naming precisely.

AI is genuinely good at spotting inconsistency, particularly with a protocol, between the protocol and supporting core study documents, between planned assumptions and historical site performance patterns, etc. It can surface operational risk signals early and identify where feasibility assumptions do not hold against real-world data. Used well, it makes what was previously invisible to a planning team visible before it becomes a problem.

Where this becomes complicated is when what AI surfaces is not an anomaly but a reflection of underlying design decisions that were never interrogated. If a protocol relies on manual workarounds, unspoken rules, and undocumented human intuition to function, it will not survive automation. The AI does not adapt to that operational culture. It will simply expose it.

I have worked on programs where cohort-level challenges traced back to exactly this dynamic. Physicians enrolling patients to provide access to a therapy but withdrawing those patients before critical endpoints were realized. Country-level regulatory friction that had roots in standard of care conflicts identifiable in the design phase but never surfaced early enough to act on. These situations are costly and they create complexity that is difficult to unwind. The situation is further exacerbated when operational pace increases.

The Data Problem Compounds This

There is a quieter problem underneath the implementation question, and it is the one I think about most. We are in the early stages of building the AI infrastructure that will shape how clinical trials run for the next decade. The training data, the historical trial information we load into these systems, the outcome patterns we teach our models to recognize as success, all of it is being assembled right now. If we build that infrastructure on top of trial records that contain operational inconsistencies, site performance gaps, and enrollment patterns shaped by protocol design flaws, we are not capturing lessons learned. We are encoding them which will result in a continuance of critical operational challenges and inconsistencies.

A trial that produced a successful regulatory outcome despite operational inefficiency does not represent a clean operational model. Loading it without that context creates bias which we want to avoid. The AI system learns that certain patterns are acceptable because the end result was a product approval. The operational friction which existed along the way disappears into the summary.

This is not a reason to slow down AI implementation. It is a reason to be deliberate about what we are building and intentional about the data we are curating as we build it.

What We Need to Prioritize Before the Next Layer Goes In

Rigorous process mapping before AI integration is not a conservative position. This is a strategic one we must be deliberate about. Before an intelligent layer is added to any part of the clinical trial workflow, the teams responsible for protocol design and operational execution need to be asking a specific set of questions.

What are the AI tools being implemented, and what is their specific operational purpose? What assumptions are embedded in how they were built, and do any of those assumptions create blind spots for the patient populations or regional operations we are working with? What are the guardrails that protect safety, privacy, fairness, and equity? And critically, where does human judgment remain the non-negotiable decision point?

Human-in-the-loop is not a phrase we should use to describe a checkbox in a validation process. It should describe how the operational team engages with AI output as a real-time thinking partner that can be overridden, challenged, and improved. This requires people who understand the domain deeply enough to know when to push back. Which means it also requires protocol designs that are sound enough that the people working with them can actually distinguish a good signal from a design flaw. If we skip this foundational work, AI may help us move faster, but faster in the wrong direction is not progress.

The Question I Am Sitting With

As an industry, we have an opportunity right now to shape what governed, responsible AI implementation looks like in clinical operations.

What would need to change in how your organization approaches protocol design for AI integration to actually deliver on its promise rather than accelerate what was already fragile?