IntuitionLabs

INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

Pharma AI Pilots: Why PoCs Fail and Scaling Strategies

> Build practical AI for pharma and biotech with IntuitionLabs. We help life-science teams turn complex information and workflows into useful software, governed knowledge systems and AI tools.

Published by IntuitionLabs · 2026-03-04 · pharma ai · ai scaling · machine learning · proof of concept · life sciences ai · data infrastructure · ai governance · pharmaceutical manufacturing

Original article: https://intuitionlabs.ai/articles/pharma-ai-pilots-failure-scaling-guide



In this article

1. Executive Summary
2. Introduction
3. Background: AI in Pharma
4. Definitions: PoC, Pilot, Production
5. The Pilot-to-Production Imperative
6. The Current State of Pharma AI Adoption
7. Industry Surveys and Statistics
8. Why Most Pharma AI Pilots Fail
9. 1. Data and Infrastructure Challenges
10. 2. Strategic and Business Alignment
11. 3. Organizational and Cultural Factors
12. 4. Technology and Integration Barriers
13. 5. Regulatory, Quality, and Compliance Hurdles
14. Case Studies and Examples
15. A Practical Scaling Playbook
16. Discussion and Future Directions
17. Conclusion

Executive Summary

Pharmaceutical companies increasingly view AI/ML technologies as strategic catalysts to accelerate research, improve manufacturing, and enhance commercial operations (moneyweek.com) (www.mckinsey.com). However, scaling generative AI and realizing financial value remain challenges. McKinsey's late-summer 2024 survey of more than 100 pharma and medtech leaders found that **32%** reported taking steps to scale generative AI, while **5%** reported competitive differentiation with consistent, significant financial value (McKinsey). These are respondent-level outcomes, not a 5% pilot-to-production conversion rate. Separately, Project NANDA's **July 2025 preliminary cross-industry generative-AI report** described limited measurable P&L impact; its evidence included interviews at 52 organizations and a survey of 153 leaders (original report). Selection bias, varying success definitions and a six-month observation period limit interpretation. These findings provide context rather than a measured pharma-wide AI project failure rate.

Pharma AI PoCs face a convergence of challenges: poor data quality and accessibility, fragmented IT/operational systems, unclear strategic alignment, missing ROI metrics, talent gaps, and stringent regulatory/compliance constraints (www.techradar.com) (www.mckinsey.com) (pharmaphorum.com). Many projects "fall in love with the technology" instead of the business outcome (www.techradar.com). Left unchecked, these problems cause a "pilot purgatory" where AI remains a novelty demo rather than a deployed solution.

This report examines barriers to scaling pharma AI pilots and proposes a **practical scaling playbook**. Building on industry studies and case examples, we outline core failure modes and evidence-based mitigation strategies. Key recommendations include: aligning AI initiatives tightly with corporate strategy and ROI metrics; building robust data infrastructure (including integration of structured and unstructured sources); establishing cross-functional governance (with executive sponsorship and AI "champions"); investing in specialized AI talent and change management; and implementing rigorous validation and compliance processes (model explainability, data security, audit trails) to satisfy regulators (www.mckinsey.com) (pharmaphorum.com). We illustrate these points with real-world examples, including Bristol Myers Squibb's Workbench clinical-trial operations platform (Accenture) and its internal talent marketplace. Finally, we discuss future directions – from federated learning and generative AI in drug discovery to evolving FDA guidance – and outline implications for pharma organizations that wish to move from hopeful pilots to productive, enterprise-scale AI deployments.

Introduction

Background: AI in Pharma

The pharmaceutical industry has long been heralded as ripe for disruption by artificial intelligence. Advances in machine learning promise to **shorten drug discovery cycles, optimize clinical trial design, and improve manufacturing efficiency** (moneyweek.com) (www.mckinsey.com). Analysts estimate that AI – especially *generative AI* – could unlock **$60–110 billion per year** in value in pharma and medical products by boosting productivity across drug discovery, development, and commercialization (www.mckinsey.com). For example, AstraZeneca's Chief Data Scientist noted that "data science and AI are transforming R&D," with applications from target identification to trial optimization (moneyweek.com). Similarly, startups like Insitro have emerged to "analyze huge datasets" of biological and chemical markers, signing partnerships with Eli Lilly and Bristol-Myers Squibb to accelerate discovery (apnews.com).

These opportunities have led to **widespread experimentation with generative AI**. In McKinsey's late-summer 2024 survey of more than 100 pharma and medtech leaders, all respondents reported experimentation; **32%** reported steps toward scaling and **5%** reported competitive differentiation with consistent, significant financial value (McKinsey). Steps toward scaling do not establish completed rollout, and financial differentiation is distinct from production deployment.

Definitions: PoC, Pilot, Production

In this context, a **Proof-of-Concept (PoC)** or **pilot** is a small-scale, time-limited AI project intended to demonstrate feasibility on a narrow problem. These pilots typically succeed at the technical level (models meet accuracy targets in isolation) but then encounter a "valley of death" when trying to embed the solution in real operations. **Production deployment** means the AI model is fully integrated into business processes or IT systems, maintained over time, and measurably improves outcomes (e.g. reduced costs, higher yields, improved patient recruitment). The **pilot-to-production gap** refers to the common situation where an AI idea passes initial proof-of-concept tests but fails to transition to sustained, scaled use.

The Pilot-to-Production Imperative

Bridging this gap is critical for pharma. A 2016 study by DiMasi, Grabowski and Hansen estimated an average out-of-pocket development cost of $1.395 billion per approved new compound, in 2013 dollars (original study). Attrition rates are high; manufacturing is tightly controlled; and commercial pressures demand efficiencies. If AI is to deliver on its hype — e.g. by reducing clinical trial timelines or optimizing scarce synthetic processes — companies must **move beyond demos**. Observers note that most AI benefits stem not from the novelty of the algorithms, but from their *integration* into workflows with clear ROI metrics (www.tomshardware.com) (www.techradar.com). Hence, understanding why pilots fail is as important as knowing how to build good models.

The Current State of Pharma AI Adoption

Industry Surveys and Statistics

- **Organizational outcomes, not pilot conversion rates**: McKinsey surveyed more than 100 pharma and medtech leaders in late summer 2024. All reported generative-AI experiments; **32%** reported steps toward scaling, and **5%** reported competitive differentiation with consistent, significant financial value (McKinsey). The survey does not establish that only 5% of pharma AI pilots reach production.
- **Cross-industry context and limitations**: Project NANDA's July 2025 report presents preliminary generative-AI research, including interviews at 52 organizations and a survey of 153 leaders, and describes limited measurable P&L impact (original report). Its authors note possible selection bias, varying success definitions and a six-month observation period that may understate longer-term success. This is contextual evidence, not a manufacturing-specific or pharma-wide pilot failure rate.
- **Investment optimism and potential value**: Despite low realized successes, companies remain optimistic. Over two-thirds of respondents in McKinsey's survey planned to *increase* AI investment (www.mckinsey.com). McKinsey projects $60–110 billion/year in latent value, especially if AI is applied strategically across discovery, development, and commercialization (www.mckinsey.com). For example, 38% of surveyed pharma leaders cited **R&D** (research) as the top area for AI, 28% cited **commercial operations** (www.mckinsey.com), reflecting that many see AI as a lever in core functions.
- **Case examples of adoption**: Several pharma leaders publicly report using AI. AstraZeneca is "applying AI throughout the discovery and development process" (moneyweek.com), and has become a major AI stock pick. Bristol Myers Squibb (BMS) and Accenture describe *Workbench* as a clinical-trial operations platform offering insights and recommendations (Accenture). BMS also piloted *MyGrowth* for talent matching; Eightfold reports reduced hiring times during its 16-week pilot (Eightfold). Insitro (a biotech AI startup) has deals with Eli Lilly and BMS to use ML on biological data (apnews.com). These examples illustrate early production use but remain exceptions rather than the rule.

In summary, the **current state** is one of broad engagement but few scaled successes. The organizational lessons from surveys and case histories consistently point to non-technical barriers as the main culprits. The following sections break down those key failure areas and discuss corresponding solutions.

Why Most Pharma AI Pilots Fail

Organizational, data and integration barriers can obstruct promising pilots. Technical feasibility also matters: RAND's practitioner research identifies limits to what AI can accomplish (RAND). Assess model performance, scientific validity and generalization alongside workflow integration and governance, especially for discovery and clinical uses. The following categories help organize that assessment.

1. Data and Infrastructure Challenges

- **Fragmented, poor-quality data**: AI thrives on data, but pharma data is often siloed. Clinical trial databases, manufacturing logs, electronic health records, and commercial metrics frequently reside in different systems with inconsistent formats. TechRadar reports for manufacturing (similarly briefing pharma) note that "the data beneath" AI models is often "fragmented, poor quality or locked away in silos," preventing reliable training or scaling (www.techradar.com). In pharma R&D, for example, laboratory data might not link easily with real-world patient data or legacy trial registries. Legacy IT systems exacerbate this: outdated databases and non-cloud architectures hinder large-scale AI (www.techradar.com).
- **Data readiness and integration**: Transforming raw data into "AI-ready" form is a monumental task. Companies must **extract, clean, integrate, and label** data from disparate sources. Without a systematic data strategy, pilots use only a small, handpicked dataset and fail when faced with real-world variability. Analysts emphasize building robust data governance and pipelines upfront. For generative AI in pharma, this includes not only structured data but also unstructured clinical notes or literature (pharmaphorum.com). In fact, Pharmaphorum recommends creating "generative AI–ready datasets (GRDs)" by curating unstructured doctor's notes, lab images, or social-media health signals as part of the pipeline (pharmaphorum.com).
- **Inadequate computing and tool integration**: Many pilots rely on vendor-provided platforms or separate models, but struggle to integrate with existing systems. Project NANDA's preliminary cross-industry generative-AI report identifies brittle workflows, lack of contextual learning and poor fit with day-to-day operations as barriers (original report). In pharma, this can mean that a model is trained on a cloud platform but never connects to the on-premise manufacturing execution system, so results cannot be actioned. Ensuring compatible, scalable tech infrastructure (e.g. secure cloud integration, interfaces to SAP/ERP, LIMS, EMR systems) is therefore critical from the start.
- **Regulatory and privacy constraints**: Pharma data often contains patient health information or proprietary research. Compliance with GDPR/HIPAA, 21 CFR Part 11, and other regs imposes strict controls. This can complicate data movement and model training. Companies may err on the side of caution and restrict data access, which starves pilots of sample size. Leaders must balance privacy with the need for data: e.g. using anonymization, federated learning, or synthetic data to comply while enabling AI. However, doing so adds layers of complexity that can derail projects without careful planning. In short: *"data is valuable but hard to share."*

Mitigation: Treat data readiness as a first-class project. Examples of best practice include establishing a unified data lake or governed data fabric, investing in data quality (cleaning and normalization), and integrating data early. As Pharmaphorum highlights, *"you need to think about every step – from acquiring data to making it useful,"* including handling unstructured sources (pharmaphorum.com). Organizations should audit their data sources, build cross-department data teams, and invest in ETL and MLOps tools **before** running pilot models. Use domain knowledge in data labeling (e.g. clinical ontologies) to ensure AI models get the right signals. In one table summarizing these issues:

Barrier	Impact on Pharma AI Pilots	Mitigation Strategies (with sources)
Fragmented/Dirty Data	Model fails to generalize; unreliable predictions; stalled training	Create integrated data lakes/warehouses; enforce data governance; curate "AI-ready" datasets (incl. unstructured) (pharmaphorum.com) (www.techradar.com); invest in data cleaning tools.
Legacy IT Systems	Incompatibilities with modern AI tools; security gaps; slow data access (www.techradar.com)	Modernize infrastructure (cloud adoption, APIs); interface pilots with enterprise systems (e.g. EMRs, MES) (www.techradar.com).
Data Privacy/Regulatory	Legal barriers to using patient/clinical data; slowed model training; compliance risk	Employ anonymization, federated learning, and secure data enclaves; build compliance into design; maintain audit logs (pharmaphorum.com).
Poor Data Integration	Inability to combine lab/clinical/sales data leads to isolated solutions; duplicative efforts	Develop ETL pipelines; use data catalogs; assign data stewards; implement master data management.

2. Strategic and Business Alignment

- **Lack of clear strategy and ROI**: A root cause of many failures is starting the pilot for technology's sake rather than solving a concrete business problem. McKinsey found about **75%** of pharma/medtech organizations lacked an enterprise-wide genAI vision or roadmap with clear success metrics (www.mckinsey.com). They often proceed in a "use case by use case" fashion without strategic guidance. Similarly, TechRadar observes organizations "fall in love with the technology, not the outcome," failing to tie projects to key objectives (www.techradar.com). Without defined ROI expectations, pilots are often abandoned as "interesting demos" that managers cannot justify continuing.
- **Misaligned priorities**: When pilots are selected by IT champions or consultants without user input, the results can be irrelevant to end users. For example, a machine-learning model to predict production yield is useless if scheduling teams cannot integrate its forecasts into their planning. Surveys show that without a strategic framework, projects proliferate as silos: "multiple pilots lead to quality, cost and sustainability challenges" and inhibit sharing of lessons (www.mckinsey.com).
- **Executive sponsorship and governance**: AI projects often lack strong C-level backing. This means insufficient funding, skimpy timelines, and little authority to make systemic changes. As Pharmaphorum advises, AI (and especially GenAI) initiatives need **C-suite sponsorship** and a governance committee to pick "winners" among pilots (pharmaphorum.com). Without this, teams may start pilots only to see them deprioritized when office politics or other crises emerge.

- **Overhyped expectations**: The surge of interest in generative AI (e.g. ChatGPT) has created unrealistic expectations for what AI can do today in pharma. Some leaders expect breakthroughs in days instead of years, or assume AI will "fix" broken processes by itself. When pilots deliver incremental gains (as is typical), the disappointment can kill further investment. This Gartner/MIT theme suggests framing AI as one component of a transformation, not a miracle.

Mitigation: Develop a **clear AI strategy and governance** before launching pilots. Define use cases in terms of business value (e.g. "reduce trial enrollment time by 20%" or "cut quality control rework"), and agree on metrics with stakeholders. Ensure alignment with corporate goals: as a Pharmaphorum piece states, start by asking "How does GenAI fit into your company's goals?" and get C-level buy-in (pharmaphorum.com). Set up an AI steering committee with representatives from IT, business, compliance, and operations to evaluate pilot proposals against this strategy. Prioritize use cases by impact and feasibility. Require pilot teams to produce business cases with expected NPV/ROI and track these in dashboards. Treat AI investment similarly to any other capital project with periodic reviews (www.techradar.com) (pharmaphorum.com).

3. Organizational and Cultural Factors

- **Skill gaps and talent shortages**: Pharma organizations often lack in-house AI expertise. Data scientists are hired for a pilot, but after initial model development the team may not have the capabilities to productize and maintain it. McKinsey notes only ~6% of firms even did a skills-based AI talent gap assessment (www.mckinsey.com). Furthermore, domain experts (e.g. clinical scientists) may be unfamiliar with AI, leading to mistrust or misaligned expectations.
- **Silos and lack of collaboration**: Traditional boundaries (R&D vs manufacturing, IT vs OT, clinical vs commercial) can impede AI projects. TechRadar emphasizes the divide between IT (information technology) and OT (operational technology) in manufacturing; similarly, pharma's drug R&D, manufacturing, and commercial areas often operate separately (www.techradar.com). Without a culture of cross-functional collaboration, data and insights do not flow, and teams may compete rather than cooperate.
- **Change management and user adoption**: Even a successful pilot that creates a working AI tool can fail if nobody uses it. Studies show that organizations must actively manage the people side of AI deployment. If front-line users (scientists, operators, sales reps) are not involved early, they may distrust the AI or find it disrupts their habits. The Pharmaphorum "playbook" emphasizes **champion users and training**: identifying target-user "champions" who advocate for the tool, providing comprehensive training, and demonstrating early wins to build confidence (pharmaphorum.com).
- **Internal resistance and incentives**: In some cases, pilots fail because their results threaten existing job roles or business models. For example, a promotion forecasting model might be ignored if marketing incentives reward locking in legacy campaigns. Or production staff may underutilize predictive-maintenance alerts to avoid scrutiny. Aligning incentives (e.g. reward teams for adopting AI-driven insights) is therefore crucial. Otherwise, pilots may produce "plugged leaks" rather than truly embedding new practices.

Mitigation: Foster an **AI-ready culture**. This includes workforce upskilling (teaching staff how to interpret and trust AI outputs), clear communication of benefits, and change management. Best practices include:

- Securing **executive sponsorship** to champion the AI project at all levels (pharmaphorum.com).
- Creating **cross-functional teams or "tiger teams"** (data scientists, domain experts, IT engineers, end-users) to co-develop solutions (pharmaphorum.com).
- Identifying **domain champions**: respected users who can provide feedback and evangelize the new tool (pharmaphorum.com).
- Running user training and pilot demonstrations early to show "tangible improvements" (pharmaphorum.com).
- Building trust through transparency and involving staff in design (so the AI augments their work, not replaces it).

4. Technology and Integration Barriers

- **Ad hoc pilot tools vs. enterprise IT**: Often, pilots rely on point solutions, prototypes, or cloud AI APIs that are easy to test but hard to deploy. Once the pilot is done, the company may discover that there is no systematic way to integrate the model into existing applications or processes. For example, an ML model may run on an offline Jupyter notebook, but there is no API link from the company's decision-support system to call the model. This disconnect means the pilot remains a standalone demo.
- **Lack of production-level software development**: Scientific prototypes often lack the robustness required for production: no automated monitoring, no user interface, no error handling. Data scientists may not follow software engineering best practices. As a result, by the time the pilot ends, the code is not production-ready. It requires rewriting, testing, and documentation, which can be as much work as the original modeling. Many organizations underestimate this "post-pilot engineering" work, leading to abandonments.
- **Evolving models and maintenance**: AI models age or "drift" as underlying data or processes change. In pharma, evolving regulations or new lab equipment can break datasets. If no plan exists for retraining or monitoring model performance, the tool will degrade in accuracy. Without dedicated MLOps pipelines and analytics monitor, pilots cannot be sustained.

Mitigation: Build pilot solutions with an eye toward production from the start. Key practices include:

- **Pilot in place**: Whenever possible, deploy the pilot within the target environment. For an operational example, BMS's Workbench provides recommendations to clinical project teams (Accenture). Design and test interfaces around the intended users and their actual decisions.
- **Agile + DevOps**: Use iterative sprints and continuous integration/continuous deployment (CI/CD) practices even in pilot phases. Automate testing, version control, and create at least a minimal user interface or API for the model.
- **Monitoring and retraining**: Include performance goals (e.g. accuracy, false positives) and set up automated monitoring. Plan for "model governance": who will refresh the model as data evolves.

- **Vendor/platform selection**: Choose enterprise-grade AI platforms that can transition from experiment to production. Ensure compatibility with security and IT policies.
- **Retrieval-augmented and explainable design**: Particularly for GenAI in pharma, combine LLMs with owned knowledge bases ("RAG") so that answers are grounded and traceable (pharmaphorum.com). This not only improves reliability but also eases validation.

5. Regulatory, Quality, and Compliance Hurdles

- **Use-specific regulatory scope**: Determine whether an AI output supports regulated decisions, affects GxP activities or falls within a medical-device framework before defining validation and change controls. FDA's January 2025 drug/biologic AI draft guidance recommends assessing model credibility for a specific context of use; it is non-binding and not for implementation. Its scope excludes drug discovery and operational-efficiency uses that do not affect patient safety, drug quality or study reliability (scope).
- **Validation and documentation**: For computerized drug-manufacturing CGMP workflows, FDA describes validation for the intended use, with its extent proportionate to system risk (drug CGMP guidance). For AI outputs supporting drug regulatory decisions, the draft credibility framework likewise tailors assessment and documentation to model risk and context. Avoid equating a pilot's transition to production with approval of a new drug.
- **Data security and privacy**: AI often requires large datasets, some of which may involve personal health data. Any processing of patient data triggers legal compliance (e.g. HIPAA, GDPR). Pilots might circumvent these by using de-identified samples; scaling up requires ironclad security to handle real patient data. As Pharmaphorum notes, in pharma *"mistakes can damage reputations and lead to legal issues"* (pharmaphorum.com). Without early involvement of risk/compliance teams, models may be disallowed in production.

Mitigation: Design AI projects with compliance in mind from inception. Implement a rigorous **model governance** framework: maintain audit trails, version histories, performance logs, and explainable outputs. Assign a "compliance officer" to each project who understands GxP rules. Engage regulators and quality assurance teams early – e.g., document data lineage and validation steps. For an ML model used in a drug-manufacturing CGMP workflow, validate the computerized workflow for its intended use, with validation proportionate to system risk, under the applicable pharmaceutical CGMP framework (FDA guidance on drug CGMP data integrity). Document controls for changes and monitoring in the pharmaceutical quality system. FDA's January 2025 drug/biologic AI draft guidance separately recommends risk-based model credibility assessment for information supporting regulatory decisions; it contains non-binding recommendations and is not for implementation. FDA's PCCP guidance concerns AI-enabled devices and should be used only within that device scope. Use model interpretability techniques where needed (feature importance, case retrieval, etc.) to assist audits. In the table below, we summarize how *Responsible AI* principles apply to pharma (data security, ethics, transparency):

Responsible AI Dimension	Pharma Implication	Recommended Practice
Data Security & Privacy	Patient data must be protected; HIPAA/GDPR apply	Use encryption, access controls; anonymize data; document data handling steps (pharmaphorum.com).
Regulatory Compliance	Assess the requirements applicable to the intended use	For computerized drug-manufacturing CGMP workflows, validate the intended use proportionately to system risk (FDA drug CGMP guidance); use the AI-device PCCP guidance only within its device scope (FDA).
Auditability & Explainability	Models need traceability for inspections; black-box decisions are risky	Keep code/artifacts in repositories; generate logs; use interpretable models when possible (pharmaphorum.com).
Ethical Use & Bias Checking	Biased AI (e.g. omitting subpopulations) can harm patients and violate law	Perform bias assessments; align with ethical guidelines; ensure fairness checks.

Case Studies and Examples

We now examine several company examples to illustrate these points:

- **Bristol-Myers Squibb (BMS) – AI in Clinical Trials**: Accenture describes *Workbench*, developed through its partnership with BMS, as a clinical-trial operations platform that combines operational data with AI-driven insights. It supports clinical project teams with recommendations and proactive trial management (Accenture case study). This example concerns operational decision support. For this playbook, the relevant lesson is to design tools around the information and decisions of their intended users.
- **Bristol-Myers Squibb – AI in Talent Management**: BMS also pursued AI for HR. In partnership with Eightfold AI, they launched *MyGrowth*, an internal talent marketplace. The system matched employees' skills to internal job opportunities. Vendor Eightfold reports that MyGrowth helped BMS shorten recruitment times (Eightfold account).
- **AstraZeneca – R&D Pipeline Integration**: AstraZeneca is known to be applying AI "throughout the discovery and development process" (moneyweek.com). While few public metrics are available, AZ's approach underscores enterprise-level commitment. Jim Weatherall of AZ notes that AI is speeding "time from science to medicine" (moneyweek.com). However, even for AZ, notable caution is expressed: the FT has observed that drug discovery using AI has not progressed as fast as hoped, largely because "crucial areas of biology… aren't sufficiently understood" (moneyweek.com). This highlights that technology must contend with scientific realities. AZ invests in both DNNs and deep domain expertise, aligning its AI strategy with biological knowledge. The lesson here is that even a well-funded leader may see incremental rather than explosive gains, and must invest in underlying knowledge (biology, process engineering) as much as in algorithms.

- **Insitro – AI Biotech Partnerships**: Insitro is an AI-based biotech founded 2018. Rather than an internal pharma project, it typifies the partnership model: Insitro uses ML to analyze large chemical/biological datasets and has signed collaboration deals with big pharma like Lilly and BMS (apnews.com). These partnerships bring AI expertise to pharma pipelines. Insitro's approach suggests one scaling lesson: dealing with data at scale (e.g. millions of *genomic* data points) and then transferring learnings to pharma collaborators. While not a *pharma company pilot*, it reflects how specialized AI vendors can help overcome the pilot-to-production gap by providing ready-made pipelines and expertise.
- **Summarized Table of Examples**: The table below summarizes these cases with outcomes and lessons:

Company/Project	Domain/Use Case	Outcome/Impact	Key Lesson
Bristol-Myers Squibb – *Workbench*	Clinical-trial operations	Accenture reports AI-driven operational insights and recommendations for trial teams (case study)	Design the platform around the information and decisions of clinical project teams.
Bristol-Myers Squibb – *MyGrowth*	Talent management (HR internal mobility)	Vendor-reported reduction in hiring times (Eightfold)	Align AI pilot with strategic goals; ensure broad user adoption (HR managers/employees as stakeholders).
AstraZeneca – Drug R&D AI	Discovery & development pipeline	Describes using AI to combine biomarkers, imaging and real-world data for patient selection (AstraZeneca)	Evaluate each use case against defined scientific and operational endpoints.
Insitro (startup)	Drug discovery (ML on biological data)	Deals with Lilly, BMS to analyze datasets (apnews.com)	Leveraging specialized AI vendors can help scale pilots by providing mature data platforms and models.

These cases illustrate different implementation settings. Accenture describes Workbench in terms of trial-execution recommendations and proactive operational decisions (case study). Public case reports illustrate possible approaches rather than establish a pharma-wide pilot success rate.

A Practical Scaling Playbook

Drawing on best practices and the above insights, we propose the following playbook for scaling Pharma AI from pilot to production. Each step is backed by industry recommendations and case lessons:

1. **Strategic Alignment & Governance**: Before piloting, define a clear AI vision. Ask, *"How does this AI project fit our business objectives?"* and secure C-level sponsorship (pharmaphorum.com). Establish an AI governance body to evaluate and prioritize projects, with agreed criteria (ROI, feasibility, risk). For example, set up a scoring rubric (expected savings vs. implementation cost). As one Pharmaphorum playbook advises, "get your C-level executives talking with your data science leaders" to ensure alignment (pharmaphorum.com). Institute stage-gates: each pilot must pass go/no-go reviews based on data readiness and pilot results.

2. **Define Metrics and ROI Upfront**: Treat AI like a capital investment: define in advance what success looks like (metrics for reduction in cost/time/error, patient outcome improvement, etc.) and track them closely (www.techradar.com). For manufacturing use-cases, this might be yield increase; for clinical use-cases, faster enrollment times; for sales, conversion uplift. Build a business case (NPV, payback) for each pilot. This combats the "pilot for publication" trap – only projects with plausible ROI proceed. If possible, run quick A/B tests or shadow trials to measure impact during the pilot.
3. **Data Strategy & Infrastructure**: Parallel to tech model building, construct an enterprise data platform. Map all relevant data sources; invest in pipelines and storage (cloud or on-prem) that can aggregate experimental and production data (pharmaphorum.com). Ensure both structured and unstructured data are captured. Use data versioning and master data management. Leverage containerization and data virtualization to facilitate secure access. Establish data governance (owners, stewards, quality metrics). This often means months of work before models train on real data – do not rush this step. Pharmaphorum's advice to "integrate and curate" data into "AI-ready datasets" (pharmaphorum.com) encapsulates this: for instance, turn unstructured physician notes into structured inputs via NLP preprocessing.
4. **Cross-Functional Team & Operating Model**: Form a dedicated AI/ML task force that spans IT, business, and domain experts (pharmaphorum.com). Each pilot should include:

- **Domain SMEs** (e.g. biologists, clinicians, engineers) to interpret results and guide labeling.
- **Data scientists/engineers** to develop and deploy models.
- **IT/DevOps staff** to integrate with company systems.
- **Regulatory/Compliance** liaisons to ensure 21 CFR/GCP requirements are met.
- **Change leads or "champions"** from the target user group to drive adoption (pharmaphorum.com). Pharmaphorum emphasizes securing executive sponsorship and using a cross-functional team as part of the operating model (pharmaphorum.com). Also, consider partnering with external AI specialists or vendors for accelerated know-how (as BMS did with Accenture on Workbench (case study) and with Eightfold) – but only if they align with your strategic goals.

5. **Agile Development & MVP Focus**: Instead of one big-bang model, start with a Minimum Viable Product that showcases value with minimal features. Use agile sprints and fast iterations. For example, if the goal is to improve yield prediction in a lyophilization process, begin by automating analysis on a single product line, not the entire facility. Keep the pilot scope narrow and measurable, then expand. This was illustrated by pharmacists who launched "early pilot projects that deliver tangible improvements to generate momentum" (pharmaphorum.com). Maintain flexibility to pivot if results are unsatisfactory or if requirements change.
6. **Robust MLOps and Tech Stack**: Invest in production-grade MLOps from the pilot stage. Use version control for data/models, automated pipelines for retraining, and monitoring dashboards. Ensure models are dockerized or containerized and can be deployed (e.g. on Kubernetes, or integrated via APIs). Most pharma firms use technologies like GPT-4 or domain LLMs for tasks like medical rep prep (pharmaphorum.com); whichever tools you choose, ensure they can handle pharma needs (e.g.

ingestion of scientific literature, secure data access). Techniques like **Retrieval-Augmented Generation (RAG)** can improve accuracy by grounding generative models on company datasets (pharmaphorum.com).

7. **User Adoption & Training**: Don't assume that a successful pilot will sell itself. Plan for **change management**: communicate the benefits in users' terms, provide hands-on training, and create feedback loops. Identify power users ("champions") in each target group who can test early versions and advocate for the tool (pharmaphorum.com). For example, train a few doctors on the AI trial-matching tool before full rollout. Collect user feedback regularly and refine the UX. Ensure help desks or support are ready when moving to production. Celebrate quick wins publicly to encourage broader acceptance.
8. **Regulatory Compliance and Responsible AI**: As deployment nears, engage quality and compliance units. Document every step: data provenance, model validation, performance testing, and decision rationale (pharmaphorum.com). Map the AI tool to existing regulatory categories (e.g. software as a medical device if applicable). Follow internal audit processes and update SOPs to include the AI tool. Implement ongoing risk management: continuously monitor for model drift, security threats, or data breaches. Incorporate explainability features (e.g. LIME/SHAP, or case examples) so that the model's recommendations can be justified to auditors. In short, operationalize **'responsible AI'** – pharmacy installations cannot afford silent errors or opacity (pharmaphorum.com).
9. **Measure and Iterate**: Once in production, measure actual impact against the original KPIs. Revisit the business case: are savings realized? Use these metrics to refine the strategy and scale further (e.g. roll out to other sites or drugs). Also collect lessons on what didn't work: did any assumptions fail? Document these in a knowledge repository to improve the next project. According to industry experts, moving from "pilot purgatory" requires capturing and sharing knowledge across initiatives (www.mckinsey.com).

These steps are summarized in the following table:

Playbook Step	Action Items	Citations/Examples
1. Strategic Alignment	Define AI vision tied to business goals. C-level sponsorship. Governance committee to prioritize pilots.	"get C-level talking with data leads" (pharmaphorum.com); 75% lacked AI strategy (www.mckinsey.com)
2. Define ROI Metrics	Specify KPIs (cost, time, quality) beforehand. Build business case (NPV/ROI). Track results scientifically.	Projects should be anchored in clear business plans (www.techradar.com)
3. Data Strategy	Develop data pipelines (structured + unstructured). Invest in integration (data lakes, ETL, catalogs). Ensure data quality and access.	Create "AI-ready datasets" from all relevant sources (pharmaphorum.com)
4. Cross-Functional Team	Assemble data scientists, SMEs, IT, compliance, and user champions. Assign clear roles (including "champions" among end-users).	"GenAI task force with cross-functional experts" (pharmaphorum.com); champion users (pharmaphorum.com)
5. Agile MVP Development	Pilot narrow-scope MVP with iterative sprints. Emphasize quick wins. Adapt scope based on feedback.	Showcase early pilots to generate momentum (pharmaphorum.com)

Playbook Step	Action Items	Citations/Examples
6. Production-grade MLOps	Use CI/CD, versioning, automated monitoring. Containerize models and integrate via APIs. Incorporate RAG and other pharma-specific AI tech.	Operational example: Workbench provides recommendations for clinical project teams (Accenture)
7. Change Management/Adoption	Train users and “champions”, communicate benefits clearly. Incorporate feedback loops. Provide ongoing support.	Champion network and training builds confidence (pharmaphorum.com)
8. Regulatory Compliance	Engage QA/reg teams; document model validation. Ensure explainability, compliance checks, audit trails.	Responsible AI with security/privacy focus (pharmaphorum.com); FDA’s AI device guidance (www.axios.com)
9. Measure & Scale	Use data-driven reviews to confirm ROI. Iterate improvements. Scale to new domains/sites. Institutionalize learnings for future projects.	Track adoption vs. projections; share knowledge to avoid silos (www.mckinsey.com)

Discussion and Future Directions

Despite the challenges, continued momentum in AI and evolving technological trends point to a promising future:

- **Generative AI and Knowledge Graphs**: Large Language Models (LLMs) and generative AI tools can accelerate literature review, hypothesis generation, and regulatory writing. Future Pharma AI may embed LLMs in lab notebooks or compliance workflows. Combining LLMs with structured domain knowledge (via “retrieval-augmented generation” (pharmaphorum.com)) can mitigate hallucinations and produce scientifically grounded answers. Early adopters are piloting AI assistants for chemists and clinicians. Over time, more subject-specific models (e.g. trained on internal corpora) will emerge.
- **Federated Learning and Privacy Tech**: To overcome data silos, federated learning enables models to train across multiple organizations’ data without sharing raw data. For instance, pharma consortia could collaboratively train on broader clinical trial data while preserving patient privacy. Advances in encryption and secure multi-party computation will make cross-company AI more feasible. The recent focus on data governance will pay off here.
- **Regulatory Evolution**: FDA’s final PCCP guidance, dated August 2025, addresses planned modifications to AI-enabled devices; it is not a general framework for pharmaceutical AI. FDA reviews these plans as part of device marketing submissions. The Pre-Cert pilot ended in September 2022. For drug and biologic uses, FDA’s January 2025 draft AI guidance describes engagement options that depend on the context of use and development program; it is non-binding and not for implementation. The EU AI Act entered into force on 1 August 2024 and became applicable on 2 August 2026, with exceptions and later application dates for certain high-risk systems. Assess each pharma or device use against the relevant system category and the current Commission timeline.
- **Digital Twins and Simulation**: Pharma manufacturing is ripe for digital twin technology (AI-driven simulations of processes). Closed-loop AI systems could self-optimize bioreactors or synthetic chemistry steps. Some companies are piloting AI-controlled “mini-factories” to reduce scale-up risk. On the R&D

side, in silico trial simulations (virtual cohorts) may speed up development. These are longer-term, but efforts now to integrate IoT and real-time analytics will lay the groundwork.

- **Workforce Transformation**: As AI becomes pervasive, pharma roles will evolve. We will see new hybrid jobs like “AI-savvy biostatistician” or “MLops engineer” in bio-manufacturing. Organizational learning will need to keep pace. Education partnerships (e.g. industry-sponsored data science academies) and rotating programs (data fellows spending time in IT, then in the lab) will help build skills.
- **Ethical and Social Implications**: Patient trust is crucial. Incidents of biased or opaque AI recommendations could hurt adoption. Pharma companies may need AI ethics boards or patient advisory input for certain applications. Transparency with the public – e.g. explaining how AI contributed to a new drug or trial design – will become part of corporate responsibility.

Conclusion

The cited McKinsey survey concerns organizations’ progress with generative AI and their financial outcomes, rather than the proportion of pharma projects that fail (McKinsey). Scaling remains an organizational challenge. By learning from past challenges and adopting a systematic, end-to-end approach, pharmaceutical organizations can build a more disciplined path from experimentation to sustained use.

Critical success factors include aligning each AI initiative with clear business value, building robust data foundations, fostering cross-functional collaboration, and rigorously addressing compliance. The playbook outlined here – gathered from industry research and real examples – provides a concrete roadmap. The examples cover distinct operational settings: Accenture describes trial-execution recommendations in BMS’s Workbench (case study), while Eightfold reports reduced hiring times during MyGrowth’s 16-week pilot (vendor account).

As AI technologies (especially generative models) and regulatory landscapes evolve, pharma stands at the cusp of a digital transformation. Future-proof companies will be those that **treat AI as a core capability**, invest in the organizational changes above, and continuously iterate. While care must be taken – given the high stakes of patient health and scientific uncertainty – the alternative is stagnation. In an industry where the next breakthrough can literally save lives, closing the pilot-to-production gap for AI is not just an operational imperative, it is a strategic mission.

References: Sources are linked inline.

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://moneyweek.com/investments/tech-stocks/how-will-ai-impact-the-pharma-industry
2. https://www.mckinsey.com/industries/life-sciences/our-insights/scaling-gen-ai-in-the-life-sciences-industry
3. https://airnd.center/v0.1_State_of_AI_in_Business_2025_Report.pdf
4. https://www.techradar.com/pro/why-ai-pilots-fail-and-how-manufacturers-can-break-the-cycle
5. https://pharmaphorum.com/digital/turning-pilots-production-how-create-scale-and-succeed-genai

6. https://www.accenture.com/en/case-studies/health/bristol-myers-squibb-accelerates-drug-development-genai
7. https://apnews.com/article/004c0ce0442b72c37bfec6e032796808
8. https://scholars.duke.edu/publication/1123333
9. https://www.tomshardware.com/tech-industry/artificial-intelligence/95-percent-of-generative-ai-implementations-in-enterprise-have-no-measurable-impact-on-p-and-l-says-mit-flawed-integration-key-reason-why-ai-projects-underperform
10. https://eightfold.ai/blog/bristol-myers-squibb-improving-employee-experience-ai
11. https://www.rand.org/pubs/research_reports/RRA2680-1.html
12. https://www.techradar.com/pro/infrastructure-modernization-is-key-to-ai-success
13. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-artificial-intelligence-support-regulatory-decision-making-drug-and-biological
14. https://www.fda.gov/media/184830/download
15. https://www.fda.gov/media/119267/download
16. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence
17. https://www.astrazeneca.com/r-d/science-and-technologies/ai-data.html
18. https://www.axios.com/2024/12/04/fda-ai-device-guidance
19. https://www.fda.gov/medical-devices/digital-health-center-excellence/digital-health-software-precertification-pre-cert-pilot-program
20. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

THE PUBLISHER

About IntuitionLabs

IntuitionLabs is an AI consulting, custom software development and data engineering firm serving pharmaceutical, biotechnology, medical-device and other life-science organizations. We work with clinical, regulatory, medical-affairs, commercial, quality and IT teams to connect technology decisions with the work people need to accomplish.

AI consulting and adoption

Our AI enablement services cover readiness assessments, use-case selection, governance and policies, team workshops, adoption measurement and ongoing advisory support. We help organizations structure the information layer behind AI: source material, context, permissions and maintained knowledge that make generated answers useful and reviewable. Private LLM inference and hosted AI options support teams evaluating how to operate AI with appropriate control over their data and infrastructure.

Software, data and life-science workflows

IntuitionLabs develops custom software for pharma and biotech, integrates enterprise systems, and builds data engineering and business intelligence solutions. Areas of focus include AI agents, regulatory research, medical writing, medical affairs, CMC information, competitive intelligence and clinical-document workflows. Our eTMF intelligence work includes cross-system reconciliation and inspection-readiness support.

Enterprise platforms and regulated delivery

We provide Veeva services, application support, managed services, integrations and custom applications, alongside enterprise content work involving platforms such as Egnyte. For regulated workflows, our services include GxP enablement, computer-system validation and software development addressing 21 CFR Part 11 requirements. The applicable controls, validation responsibilities and acceptance criteria are defined for each engagement.

Work with IntuitionLabs

Explore AI enablement, pharma and biotech software development, data engineering and BI, and Veeva services. Contact IntuitionLabs to discuss your workflow, information sources and implementation needs.

IntuitionLabs publishes educational research to help life-science teams make informed technology decisions. Coverage of a product or organization does not imply a client relationship, endorsement or partnership.

Website: https://intuitionlabs.ai

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.