McKinsey & Company

Life Sciences Practice

Boosting biopharma R&D performance with a next-generation technology stack

A modern, well-designed tech stack can unlock the potential of AI, automation, and data and their promised benefits to R&D productivity.

by Jeffrey Lewis, Joachim Bleys, and Ralf Raschke
with Moritz Wolf

January 2025

Amid growing complexities, escalating costs, and a heightened demand for innovative therapies, biopharmaceutical R&D organizations are under pressure to find ways to boost their productivity. Preclinical studies are increasingly using diverse assays, complex “omics” data, and humanized models to evaluate drug candidates. Clinical trials are adopting new trial designs and incorporating personalized-medicine approaches and even more end points.[1]

In this challenging environment, pharma leaders are turning to advanced technologies such as digital and analytics (DnA) solutions to reshape R&D operations and deliver new efficiencies, such as faster discovery of new molecular entities and novel therapeutic mechanisms. Leading companies are already reaping benefits from enhanced DnA capabilities, such as in silico modeling to accelerate molecule discovery, AI-powered tools for optimized clinical trials (including site selection, patient recruitment, and trial monitoring), and automated generative AI (gen AI) to produce higher-quality documents, which shrinks submission timelines. Gen AI tools alone are projected to unlock as much as $53 billion in annual value across the R&D value chain.

Despite these advancements, and even though US Food and Drug Administration approvals for novel therapeutics were the second-highest in the past 30 years,[2] the industry’s overall R&D productivity has been mostly flat. Development timelines still exceed a decade, R&D costs per asset on average surpass $2 billion, and barely 13 percent of assets that enter Phase 1 trials make it to launch. There are several root causes for the limited progress in realizing the promise of DnA in pharma compared with other industries. Typical failure modes we see include companies clinging to legacy processes that undermine standardization, neglecting change management, and implementing technology without clear business benefits. Additional pitfalls include conducting transformations in departmental silos while relying on inflexible systems plagued by low-quality, siloed data.

These challenges can be overcome with a well-designed, modern R&D technology stack—the IT infrastructure that enables the insights, workflows, data collection, storage, transfer, and processing of data throughout the discovery, research, and clinical-development stages. In this article—which details the sixth ingredient (out of eight) in our recipe for sustained R&D excellence introduced in our flagship *RewiR&D* article—“Making more medicines that matter”—we present the architecture for a modern tech stack for biopharma R&D organizations and discuss its role in enabling next-generation (next-gen) data, analytics, and technology.

Core layers of a modern R&D tech stack

When properly integrated and implemented, a modern tech stack powers the seamless flow and analysis of research and clinical data. Many R&D organizations rely on decentralized tech stacks that consist of independent core systems and applications connected via point-to-point interfaces. While this decentralized architecture may be simpler to implement initially, it becomes unwieldy when the demand for scaling arises.

[1] *Impact report*, Tufts Center for the Study of Drug Development, Tufts University, May/June 2021, Volume 23, Number 3; see also Charlie Passut, “Trial complexity, endpoints continue to increase, stretching site resources,” CenterWatch, January 18, 2021.
[2] *Research and development in the pharmaceutical industry*, US Congressional Budget Office, April 2021.

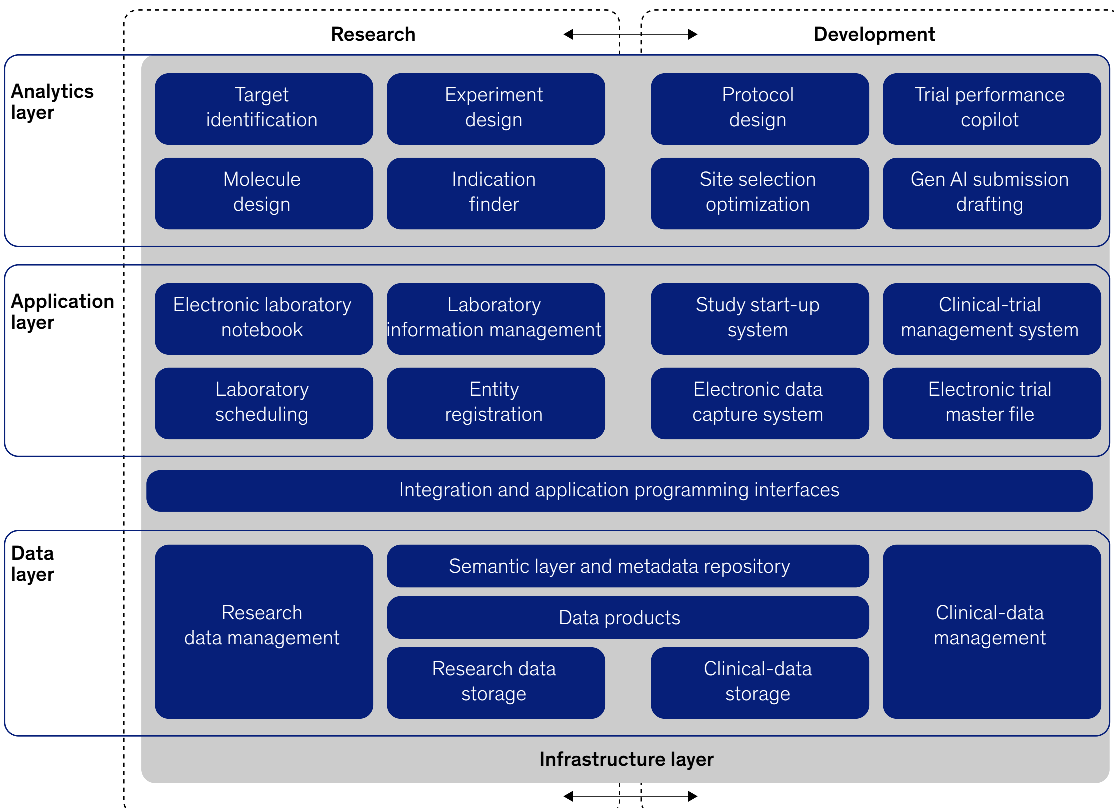
A modern tech stack can be visualized and organized as four distinct but integrated layers (Exhibit 1). This architecture allows organizations to develop and implement new workflows without being limited by the time-consuming process of updating legacy systems, facilitates the reuse of data and the implementation of AI workflows, and improves interoperability through more efficient management of data and services within and beyond the R&D functions. As a result, R&D organizations can maintain flexibility and speed in their technological development while still making the most of their existing resources.

The infrastructure layer forms the foundation, providing cloud-based resources that ensure high performance, scalability, and security for the entire system. The data layer manages the integration, curation, and accessibility of clinical and operational data, adhering to

Exhibit 1

The R&D tech stack comprises four interconnected layers, each containing the critical electronic systems used in drug discovery and development.

A drug discovery and development tech stack



McKinsey & Company

FAIR principles (findable, accessible, interoperable, and reusable).[3] The application layer comprises core clinical systems, including electronic data capture (EDC) and clinical-data management systems (CDMSs), as well as specialized tools such as electronic lab notebooks.

These applications feature integrated workflows tailored to meet the specific needs of clinical development. At the apex of the tech stack is the analytics layer, which features basic and advanced tools (such as AI and gen AI) that perform data analysis and visualization, ultimately enabling real-time insights and informed decision making throughout the R&D process.

Unlocking value from a rewired tech stack

A modern R&D tech stack can improve pharma R&D performance largely by fortifying the IT infrastructure, streamlining quality and compliance activities, and ultimately boosting R&D productivity (Exhibit 2). Here, we describe the role of the tech stack in boosting these performance factors.

Exhibit 2

A modern tech stack can unlock value across several key factors of R&D productivity.

Relevance of R&D tech stack layer maturity for value creation, rating (low = 1, high = 4)

		Cycle time	Cost	Probability of success	Innovation	Future readiness	Quality and compliance
Analytics	Research	3	1	4	3	4	1
	Development	2	1	2	2	3	2
Application	Research	2	2	1	2	1	2
	Development	3	3	1	2	1	3
Data	Research	2	2	2	3	3	4
	Development	3	3	2	3	3	4
Infrastructure	Research	1	2	1	2	1	2
	Development	1	2	1	2	1	2

HIGH

LOW

McKinsey & Company

[3] Mark D. Wilkinson et al., "The FAIR Guiding Principles for scientific data management and stewardship," *Scientific Data*, March 15, 2016.

Cycle time acceleration

A modern R&D tech stack and its analytics, application, and data layers play a critical role in accelerating every phase of the journey, from drug discovery to market launch. In the analytics layer, AI and machine learning (AI/ML) tools can be used to predict drug-target interactions, optimize trials, and enhance decision making, such as prioritization of experiments and trial issue resolution approaches. AstraZeneca's augmented drug design platform uses AI to streamline the design cycle.[4] By integrating molecule ideation, activity prediction, and synthesis prediction, AstraZeneca has reduced the number of design-make-test-analyze (DMTA) cycles required, speeding up discovery. Predictive AI models provide real-time insights into the entire asset portfolio, enabling faster decision making during trials, such as when to set up "rescue sites" to prevent delays.

In the application layer, where operational and patient data are collected, companies have begun to use gen AI at scale to create clinical-study reports and submit dossiers in record times. These systems also play a role in cycle time acceleration through automation. For example, processes associated with study start-up, clinical-trial management systems (CTMSs), EDC, and CDMSs can all be streamlined with AI and automation. On top of a modern data layer, CDMS and EDC applications can substantially reduce the time between last patient, last visit (LPLV) and database lock, enabling faster decision making and submissions.

Cost reduction

Technology is crucial for reducing high R&D costs. The analytics, application, and data layers work together to improve workflow efficiency and maximize productivity. AI and in silico modeling help prioritize experiments, cutting down on costly lab work. Automated lab workflows, facilitated by laboratory information management systems and electronic lab notebooks, streamline data handling and further reduce manual tasks. Trial management can be automated by CTMSs and EDC, which promote optimal site selection and patient recruitment to minimize delays and trial failures. The infrastructure layer can add to cost savings by providing scalable cloud resources, eliminating the need for expensive on-premises hardware.

Probability of success

Some pharma companies are beginning to see improved success rates due to the integration of advanced tools at the analytics layer. In the research phase, biotech companies—such as Recursion's Operating System—are using AI-powered hypotheses to help them select millions of experiments that can run weekly with a combination of in silico modeling and wet-lab automation. In the area of clinical development, researchers use AI to analyze a combination of real-world data and clinical-trial data, which can help them improve trial design and patient recruitment. Gen AI is also being used to produce higher-quality regulatory submissions, which can boost the odds for successful first-cycle approvals.

Key to these improvements is the modern tech stack architecture that operates across all four critical layers. Integration of the data layer and infrastructure layer enables analysis of multiple internal and external data sets, such as historical records, real-world evidence (RWE), external publications, and genome-wide association studies. Insights can then be drawn from those analyses to help researchers refine predictions, determine optimal experiments, and make more informed decisions about which targets to pursue, ultimately driving up the probability of success.

[4] "Bio-IT World names 2022 Innovative Practices Awards winners," Bio-IT World, April 11, 2022.

Innovation

Innovation is crucial for addressing unmet medical needs and refreshing biopharma portfolios. Here, too, the tech stack's layers play pivotal roles. The analytics layer propels discovery by simulating biology and disease, identifying molecular targets, and designing promising drug candidates that proceed through lab tests to clinical trials. The data layer manages, integrates, and supplies diverse data—such as omics data and RWE—to continually train and improve models.

Some pharma companies are also exploring ways to share data while protecting intellectual property. Federated learning is one approach, allowing companies to contribute research findings while accessing enhanced discovery algorithms. Proprietary data on molecules and compounds remains secure through a "sanitization" process, such as is done within the European Union's MELLODDY (machine learning ledger orchestration for drug discovery) consortium.

Future readiness

Modernizing the R&D tech stack makes the pharma IT infrastructure more adaptable, enabling faster, more flexible responses to changes, such as regulatory data localization requirements. It can also allow for rapid adoption of innovations such as decentralized clinical-trial technology during the COVID-19 pandemic or, eventually, quantum computing. Streamlined tech stacks also minimize tech debt—eliminating short-term fixes of complex point-to-point integrations that often need to be replaced by APIs.

By consolidating software and migrating to the cloud, companies can reduce outdated technology, freeing IT resources to innovate with advanced automation and gen AI. Cloud infrastructure offers the scalability to invest in strategic tools, such as AI-driven decision making and automated workflows. In our experience, these shifts can free up to 30 percent of R&D IT spending, accelerating digital and AI priorities and enabling large-scale digital transformation.

Quality and compliance

In the highly regulated pharmaceutical industry, ensuring data integrity and regulatory compliance is paramount. The application and data layers promote transparency and compliance throughout the R&D process. In research, standardized data sets help teams collaborate seamlessly while maintaining data accuracy. In development, systems such as electronic clinical-outcome assessment and EDC automate data collection and report generation, ensuring compliance with global regulations. AI tools can catch compliance issues early, giving teams the chance to address them proactively. And real-time insights from the analytics layer, by finding and fixing trial issues early, can result in higher-quality approval submissions and faster regulatory approvals.

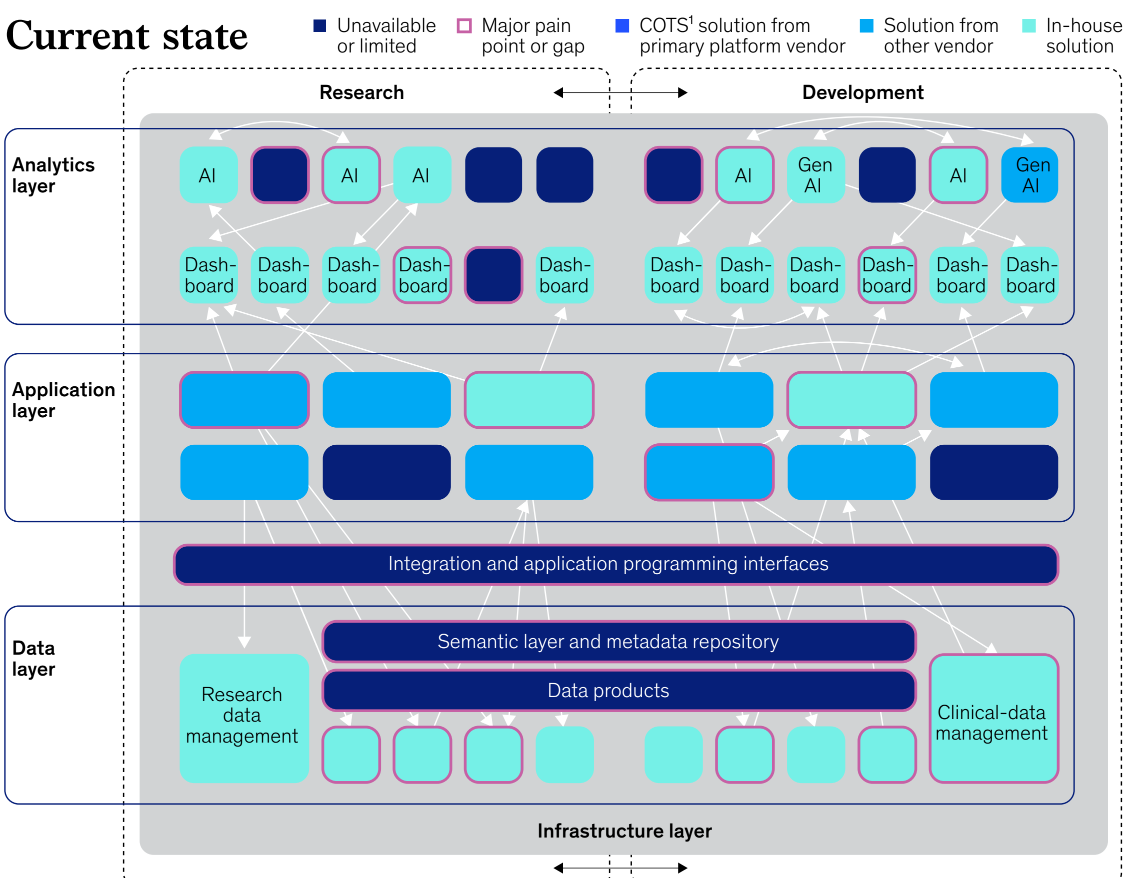
Designing the modern R&D tech stack

Each biopharma organization should tailor its tech stack design to its needs. However, a few general principles exist for a modern tech stack that set it apart from the legacy designs (Exhibit 3).

A modern pharma R&D tech stack should be built on a modular, flexible architecture that can support software-as-a-service (SaaS) platforms for core applications, cloud-based platforms for data management, and advanced DnA for differentiating insights. The modularity enables pharmaceutical companies to adopt and integrate new technologies as they emerge. Also, cybersecurity protocols such as end-to-end encryption and role-based access controls can be embedded in each layer to protect sensitive data and achieve compliance with the Health Insurance Portability and Accountability Act in the United States, the European Union's General Data Protection Regulation, and other security regulations. Overall, the modern tech stack architecture is simplified through standardizing common platforms and reducing the number of point solutions, integrations, and ad hoc workarounds.

Exhibit 3

The typical current-state R&D technology stack can be transformed for value generation through seven shifts.

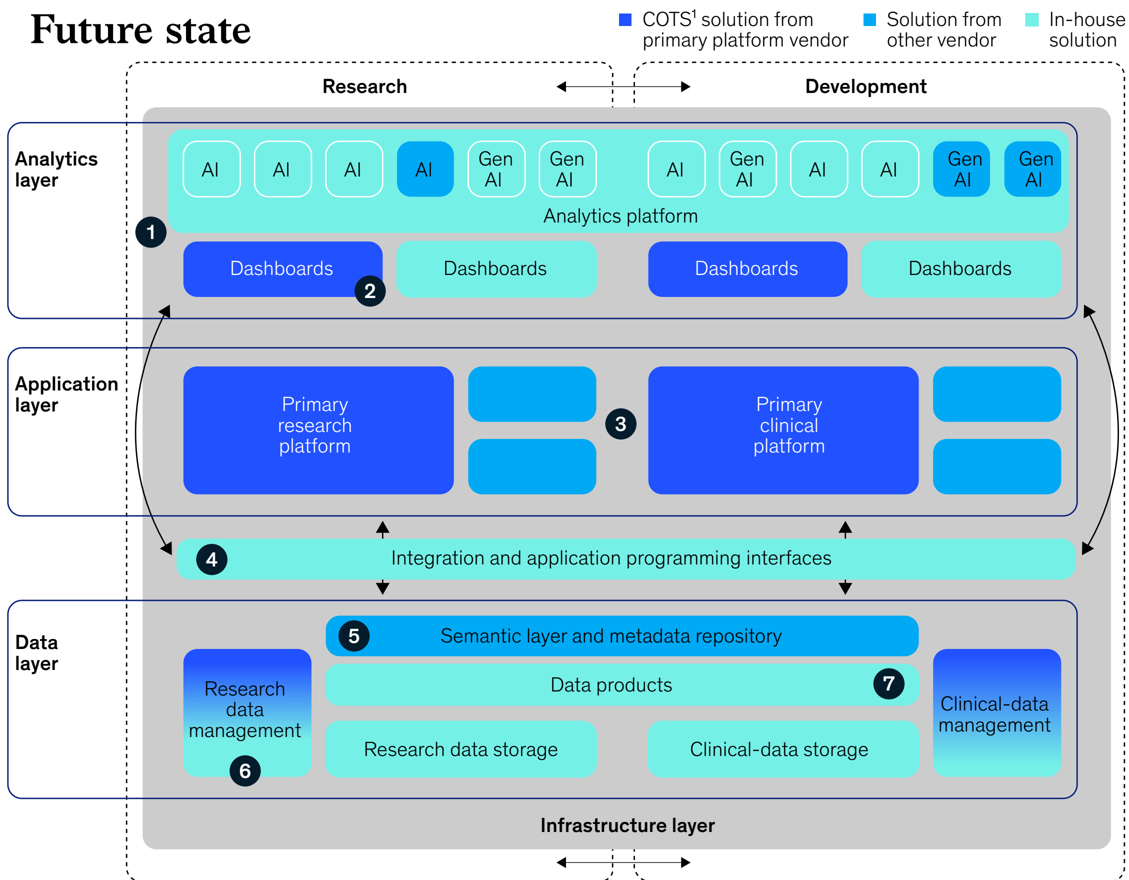


[1]Commercial, off the shelf.

McKinsey & Company

Exhibit 3 (continued)

The typical current-state R&D technology stack can be transformed for value generation through seven shifts.



1. An integrated analytics platform can consolidate, and accelerate scaling of, scattered AI pilots, enabling standardized AI deployments and component reuse.

2. Modern SaaS platforms capable of real-time tracking can replace multiple customized dashboards, improving user experiences for such applications such as clinical trial management systems and enabling resources to be shifted from maintenance to innovation.

3. Modular SaaS platforms can replace fragmented customized applications, enabling intuitive, standardized solutions with automated workflows and real-time collaboration for areas such as lab and clinical operations, safety, and regulatory.

4. A centralized API-based data exchange can replace brittle point-to-point integrations, connecting tech stack layers, use cases, systems, and data sources.

5. Metadata-driven digital data flow can replace manually transcribed, document-centric data flow, enabling automated processes such as the generation of case report forms and clinical study reports.

6. Automated real-time data cleaning, query management, and transformation can replace manual data management, enabling faster cycle times and improved data quality.

7. Reusable data products that consistently feed multiple AI and machine learning models can replace manual data engineering for secondary data use, enabling greater efficiency for such processes as trial site selection.

[1]Commercial, off the shelf.

McKinsey & Company

Analytics layer
This layer serves as the system's brain, using AI/ML to provide insights into drug efficacy, patient outcomes, and trial optimization. APIs enable researchers to access reusable AI models and data analysis tools and apply solutions quickly across various projects. For example, this configuration could allow real-time analysis of data from wearable devices worn during clinical trials, which could lead to more rapid protocol adjustments. The modern design would feature dashboards for routine reporting and analytics, freeing up R&D teams to focus on building the differentiating analytics and advanced insights that set their research apart.

Application layer
Most R&D capabilities at the application layer are provided by SaaS platform vendors, including software for laboratory information management systems (LIMSs) for research and CTMS software for development. Ideally, R&D organizations should seek an integrated out-of-the-box core platform that should require little configuration and customization—this could reduce costs and streamline operations.

Beyond the core, solutions from other SaaS vendors could be plugged in to address capabilities of strategic importance or fill existing gaps. Once the integrated platform is in place, APIs can ensure seamless connections between core systems such as electronic laboratory notebooks and LIMSs in research or CTMSs and EDC in development, the other tech stack layers, and third parties such as clinical research organizations, ensuring seamless data flow throughout the R&D process.

Data layer
The data layer centralizes R&D data using cloud-based platforms, such as data lakes or distributed data meshes, that facilitate efficient management of genomic and molecular information and other large data sets. Cloud platforms ensure easy accessibility for various teams and support AI-driven insights into drug development while maintaining regulatory compliance. APIs integrate disparate data sets across departments, streamlining data access and reducing analysis time.

Infrastructure layer
The infrastructure layer forms the foundation for all research processes. A hybrid cloud model maximizes both security and scalability. The hybrid approach safeguards sensitive data in secure, private environments, while less sensitive workloads reside in the public cloud. Additionally, the infrastructure layer supports "infrastructure as code," which uses automation for the creation of tasks such as molecular simulations or drug interaction analyses. This ensures agility and repeatability in research workflows.

Implementing the next-gen R&D tech stack

Building a modern R&D tech stack requires a strategic approach beyond simply adopting new tools. It begins with a thorough assessment of business needs and current technology to define a clear future vision. This process involves evaluating existing capabilities, identifying gaps, and prioritizing areas for modernization. Organizations should look to best-in-class examples and tailor a North Star vision that is specific to their needs.

To build a cutting-edge R&D tech stack, pharma leaders should address four key considerations with respect to their strategic priorities:

1. *Scope of modernization.* Leaders should first define the scope of their modernization efforts, determining which layers of the tech stack require modernization based on current maturity and gaps. They should consider their focus area (clinical development or broader R&D) and align with data strategy and value delivery goals. They should also identify areas for differentiation, particularly in the implementation of AI and data analytics—that exercise could yield high-value use cases such as AI-powered indication finding or patient dropout predictions.

2. *Modernization archetype.* Leaders should consider which tech stack architecture is most aligned with their organization's strategic priorities: platform, best of breed, or hybrid. A platform approach implements ready-to-use solutions from a few vendors, complemented by bespoke additions. Best-of-breed solutions involve customizing various vendors' offerings for specific use cases. And a hybrid architecture merges platform solutions with best-of-breed tools in specific areas for differentiation.

3. *Vendor combinations.* Selecting the optimal combination of vendors to meet R&D needs is a critical decision for leaders. This decision should prioritize seamless interoperability across tech stack layers, strike a balance between user-friendliness and quality, and mitigate vendor lock-in risks.

4. *Ownership and collaboration.* Successful application modernization requires a synergistic partnership between business and IT teams, with business goals providing the strategic direction and IT teams facilitating the implementation journey. IT should act as a strategic enabler, translating overarching business goals into well-defined technical requirements and actionable use cases. Heads of R&D and chief information officers should collaborate closely to ensure the seamless adoption and integration of the new system throughout the organization.

Biopharma leaders can take a strategic approach and establish a next-gen R&D tech stack that enhances efficiency and enables the adoption of advanced technologies, ultimately leading to improved R&D performance and the potential for significant value capture. A flexible, modular IT infrastructure helps pharmaceutical companies stay competitive, adapt to new challenges, and accelerate the development of lifesaving treatments for patients.

Jeffrey Lewis is a senior partner in McKinsey's New Jersey office, where **Ralf Raschke** is a senior expert; **Joachim Bleys** is a senior partner in the Carolinas office; and **Moritz Wolf** is an associate partner in the Berlin office.

The authors wish to thank Jon Boorstein, Julius Seitter, Svenja Steinbach, and Yevgen Kanalosh for their contributions to this article.

This article was edited by Jermey Matthews, an editor in the Boston office.

Designed by McKinsey Global Publishing