

April 2013

# How big data can revolutionize pharmaceutical R&D

**Jamie Cattell, Sastry Chilukuri, and Michael Levy**

Pharmaceutical R&D suffers from declining success rates and a stagnant pipeline. Big data and the analytics that go with it could be a key element of the cure.

**After transforming** customer-facing functions such as sales and marketing, big data is extending its reach to other parts of the enterprise. In research and development, for example, big data and analytics are being adopted across industries, including pharmaceuticals.

The McKinsey Global Institute<sup>1</sup> estimates that applying big-data strategies to better inform decision making could generate up to \$100 billion in value annually across the US health-care system, by optimizing innovation, improving the efficiency of research and clinical trials, and building new tools for physicians, consumers, insurers, and regulators to meet the promise of more individualized approaches.

The big-data opportunity is especially compelling in complex business environments experiencing an explosion in the types and volumes of available data. In the health-care and pharmaceutical industries, data growth is generated from several sources, including the R&D process itself, retailers, patients, and caregivers. Effectively utilizing these data will help pharmaceutical companies better identify new potential drug candidates and develop them into effective, approved and reimbursed medicines more quickly.

## **Imagine a future where the following is possible:**

- Predictive modeling of biological processes and drugs becomes significantly more sophisticated and widespread. By leveraging the diversity of available molecular and clinical data, predictive modeling could help identify new potential-candidate molecules with a high probability of being successfully developed into drugs that act on biological targets safely and effectively.
- Patients are identified to enroll in clinical trials based on more sources—for example, social media—than doctors' visits. Furthermore, the criteria for including patients in a trial could take significantly more factors (for instance, genetic information) into account to target specific populations, thereby enabling trials that are smaller, shorter, less expensive, and more powerful.

<sup>1</sup> Our work builds on insights from several reports and articles, all available on mckinsey.com: James Manyika et al., *Big data: The next frontier for innovation, competition, and productivity*, McKinsey Global Institute, May 2011; Peter Groves et al., *The 'big data' revolution in healthcare: Accelerating value and innovation*, January 2013; and Ajay Dhankhar et al., "Escaping the sword of Damocles: Toward a new future for pharmaceutical R&D," *McKinsey Perspectives on Drug and Device R&D 2012*.

- Trials are monitored in real time to rapidly identify safety or operational signals requiring action to avoid significant and potentially costly issues such as adverse events<sup>2</sup> and unnecessary delays.
- Instead of rigid data silos that are difficult to exploit, data are captured electronically and flow easily between functions, for example, discovery and clinical development, as well as to external partners, for instance, physicians and contract research organizations (CROs). This easy flow is essential for powering the real-time and predictive analytics that generate business value.

That's the vision. However, many pharmaceutical companies are wary about investing significantly in improving big-data analytical capabilities, partly because there are few examples of peers creating a lot of value from it. However, we believe investment and value creation will grow. The road ahead is indeed challenging, but the big-data opportunity in pharmaceutical R&D is real, and the rewards will be great for companies that succeed.

### **The big-data prescription for pharmaceutical R&D**

Our research suggests that by implementing eight technology-enabled measures, pharmaceutical companies can expand the data they collect and improve their approach to managing and analyzing these data.

#### **Integrate all data**

Having data that are consistent, reliable, and well linked is one of the biggest challenges facing pharmaceutical R&D. The ability to manage and integrate data generated at all stages of the value chain, from discovery to real-world use after regulatory approval, is a fundamental requirement to allow companies to derive maximum benefit from the technology trends. Data are the foundation upon which the value-adding analytics are built. Effective end-to-end data integration establishes an authoritative source for all pieces of information and accurately links disparate data regardless of the source—be it internal or external, proprietary or publicly available. Data integration also enables comprehensive searches for subsets of data based on the linkages established rather than on the information itself. “Smart” algorithms linking laboratory and clinical data, for example, could create automatic reports that identify related applications or compounds and raise red flags concerning safety or efficacy.

Implementing end-to-end data integration requires a number of capabilities, including trusted sources of data and documents, the ability to establish cross-linkages between elements, robust quality assurance, workflow management, and role-based access to ensure that specific data elements are visible only to those who are authorized to see it. Pharmaceutical companies generally avoid overhauling their entire data-integration system at once because of the logistical challenges and costs involved, although at least one

<sup>2</sup> Adverse events refer to harm to or death of trial participants.

global pharmaceutical enterprise has employed a “big bang” approach to remaking its clinical IT systems.

Companies typically employ a two-step approach: first, they prioritize the specific data types to address (usually clinical data) and create additional data-warehousing capabilities as needed. The goal is to tackle the most important data first to obtain benefits as soon as possible. This step alone can take over a year and requires significant infrastructure and procedural changes. Second, the company develops an approach for the next levels of priority data, including scenario analysis, ownership, and expected costs and timelines.

#### Collaborate internally and externally

Pharmaceutical R&D has been a secretive activity conducted within the confines of the R&D department, with little internal and external collaboration. By breaking the silos that separate internal functions and enhancing collaboration with external partners, pharmaceutical companies can extend their knowledge and data networks.

Whereas end-to-end integration aims to improve the linking of data elements, the goal of collaboration is to enhance the linkages among all stakeholders in drug research, development, commercialization, and delivery.

Maximizing internal collaboration requires improved linkages among different functions, such as discovery, clinical development, and medical affairs. This can lead to insights across the portfolio, including clinical identification and research follow-up on potential opportunities in translational medicine<sup>3</sup> or identification of personalized-medicine opportunities through the combination of biomarkers research and clinical outcomes; predictive sciences could also recommend options at the research stage based on clinical data or simulations.

External collaborations are those between the company and stakeholders outside its four walls, including academic researchers, CROs, providers, and payors. Several examples show how effective external collaboration can broaden capabilities and insights:

- External partners, such as CROs, can quickly add or scale up internal capabilities and provide access to expertise in, for example, best-in-class management of clinical studies.
- Academic collaborators can share insights from the latest scientific breakthroughs and make a wealth of external innovation available. Examples include Eli Lilly’s Phenotypic Drug Discovery Initiative, which enables external researchers to submit their compounds for screening using Lilly’s proprietary tools and data to identify whether the compound is a potential drug candidate. Participation in the screening does not require the researcher to give up intellectual property, but it does offer Lilly a first look at new compounds, as well as an avenue to reach researchers who are not typical drug-discovery scientists.

<sup>3</sup>Translational medicine refers to taking new scientific discoveries and turning those into effective health improvements and medicines.

- Collaborative “open space” initiatives can enable experts to address specific questions or share insights. Examples include the X PRIZE, which provides financial incentives for teams that successfully meet a big challenge (such as enabling low-cost manned space flight), and InnoCentive, which offers financial incentives for individuals or teams that address a specific problem (such as determining a compound’s synthesis pathway).
- Customer insights can be used to shape strategy throughout the pipeline progression.

Some pharmaceutical companies have made inroads in improving internal and external collaboration, which involves addressing a number of challenges. These include putting in place communications systems and governance to enable appropriate and effective information exchange. Another challenge is to promote a shift in mind-set, moving away from withholding all data and toward identifying which data can be shared and with whom. In addition, pharmaceutical enterprises must understand and mitigate the legal, regulatory, and intellectual-property risks associated with a more collaborative approach.

Some pharmaceutical companies start to improve collaboration by identifying data elements to share with specific sets of trusted partners, such as CROs, and establishing privileged and near-real-time access to data produced by external partners. Such steps are only the beginning, however, as they are essentially just a way to expand the “circle of trust” to select partners.

#### Employ IT-enabled portfolio-decision support

To ensure the appropriate allocation of scarce R&D funds, it is critical to enable expedited decision making for portfolio and pipeline progression. Pharmaceutical companies often find it challenging to make appropriate decisions about which assets to pursue or, sometimes more important, which assets to kill. The personnel or financial investments they have already made may influence decisions at the expense of merit, and they often lack appropriate decision-support tools to facilitate making tough calls.

IT-enabled portfolio management allows data-driven decisions to be made quickly and seamlessly. Smart visual dashboards should be used whenever possible to allow rapid and effective decision making, including for the analysis of current projects, business-development opportunities, forecasting, and competitive information. These visual systems should provide high-level dashboards that permit users to deeply examine the data, including information to bolster managerial decision making as well as detailed tactical information, and that make asset performance and opportunities more transparent.

In addition to the technical requirements, portfolio decision making should follow a defined process with known timing, deliverables, service levels, and stakeholders. The people involved in the process should be given clear roles and authority (for example, their ability to make decisions should be defined). Resource allocation should be based on a

systematic approach that accommodates top-down budgetary requirements and bottom-up requests. And innovation boards at the corporate level and at the business-unit or therapeutic-area level should review the portfolio regularly. The boards should assess, manage, and prioritize the portfolio based on the corporate strategy and changes in the business landscape or industry context.

#### Leverage new discovery technologies

Pharmaceutical R&D must continue to use cutting-edge tools. These include sophisticated modeling techniques such as systems biology and high-throughput data-production technologies—that is, technologies that produce a lot of data quickly, for example, next-generation sequencing, which, within 18 to 24 months, will make it possible to sequence an entire human genome at a cost of roughly \$100.

The wealth of new data and improved analytical techniques will enhance future innovation and feed the drug-development pipeline.

Integrating vast amounts of new data will test a pharmaceutical company's analytical capabilities. For example, a company will need to connect patient genotypes to clinical-trial results to identify opportunities for improving the identification of responsive patients. Such developments would make personalized medicine and diagnostics an integral part of the drug-development process rather than an afterthought and would lead to new discovery technologies and analytical techniques.

#### Deploy sensors and devices

Advances in instrumentation through miniaturized biosensors and the evolution in smartphones and their apps are resulting in increasingly sophisticated health-measurement devices. Pharmaceutical companies can deploy smart devices to gather large quantities of real-world data not previously available to scientists. Remote monitoring of patients through sensors and devices represents an immense opportunity. This kind of data could be used to facilitate R&D, analyze drug efficacy, enhance future drug sales, and create new economic models that combine the provision of drugs and services.

Remote-monitoring devices can also add value by increasing patients' adherence to their prescriptions. Examples of devices that are under development include smart pills that can release drugs and relay patient data, as well as smart bottles that help track usage. Technology and mobile providers are offering services such as data feeds, tracking, and analysis to complement medical devices. These devices and services, combined with in-home visits, have the potential to decrease health-care costs through shortened hospital stays and earlier identification of health issues.

### Raise clinical-trial efficiency

A combination of new, smarter devices and fluid data exchange will enable improvements in clinical-trial design and outcomes as well as greater efficiency. Clinical trials will become increasingly adaptable to react to drug-safety signals seen only in small but identifiable subpopulations of patients. Examples of potential clinical-trial efficiency gains include the following:

- Dynamic sample-size estimation (or reestimation) and other protocol changes could enable rapid responses to emerging insights from the clinical data. Efficiency gains are achieved by enabling smaller trials for equivalent power or shortening the time necessary to expand a trial.
- Adapting to differences in site patient-recruitment rates would allow a pharmaceutical company to address lagging sites, bring new sites online if necessary, and increase recruiting from more successful sites.
- Increased use of electronic data capture could help in recording patient information in the provider's electronic medical records. Using electronic medical records as the primary source for clinical-trial data rather than having a separate system could accelerate trials and reduce the likelihood of data errors caused by manual or duplicate entry.
- Next-generation remote monitoring of sites, enabled by fluid, real-time data access, could improve management and responses to issues that arise in trials.

### Improve safety and risk management

Pharmaceutical companies can use safety as a competitive advantage in regulatory submissions and after regulatory approval, once the drug is on the market. Safety monitoring is moving beyond traditional approaches to sophisticated methods that identify possible safety signals arising from rare adverse events. Furthermore, signals could be detected from a range of sources, for example, patient inquiries on Web sites and search engines. Online physician communities, electronic medical records, and consumer-generated media are also potential sources of early signals regarding safety issues and can provide data on the reach and reputation of different medicines. Bayesian analytical methods, which can identify adverse events from incoming data, could highlight rare or ambiguous safety signals with greater accuracy and speed.

An early response to physician and patient sentiments could prevent regulatory and public-relations backlashes. The FDA is investing in the evaluation of electronic health records through the Sentinel Initiative, a legally mandated electronic-surveillance system that

links and analyzes health-care data from multiple sources. As part of this system, the FDA now has secured access to data concerning more than 120 million patients nationwide.

#### Sharpen focus on real-world evidence

Real-world outcomes are becoming more important to pharmaceutical companies as payors increasingly impose value-based pricing. These companies should respond to this cost-benefit pressure by pursuing drugs for which they can show differentiation through real-world outcomes, such as therapies targeted at specific patient populations. In addition, the FDA and other government organizations have created incentives for research on health economics and outcomes.

To expand their data beyond clinical trials, some leading pharmaceutical companies are creating proprietary data networks to gather, analyze, share, and respond to real-world outcomes and claims data. Partnerships with payors, providers, and other institutions are critical to these efforts.

#### The challenges of a big-data transformation

For a big-data transformation in pharmaceutical R&D to succeed, executives must overcome several challenges.

##### Organization

Organizational silos result in data silos. Functions typically have responsibility for their systems and the data they contain. Adopting a data-centric view, with a clear owner for each data type across functional silos and through the data life cycle, will greatly facilitate the ability to use and share data. The expertise gained by the data owner will be invaluable when developing ways to use existing information or to integrate internal and external data. Furthermore, having a single owner will enhance accountability for data quality. These organizational changes will be possible only if a company's leadership understands the potential long-term value that can be unlocked through better use of internal and external data.

##### Technology and analytics

Pharmaceutical companies are now saddled with legacy systems containing heterogeneous and disparate data. Increasing the ability to share data requires rationalizing and connecting these systems. There's also a shortage of people equipped to develop the technology and analytics needed to extract maximum value from the existing data.

##### Mind-sets

Many pharmaceutical companies believe that unless they identify an ideal future state, there is little value to investing in improving big-data analytical capabilities. Indeed,

they seem to fear being the first mover, since there are few examples of pharmaceutical companies creating a lot of value from the improved use of big data. Compounding their hesitation is concern about increasing interactions with regulators if they pursue a big-data change program. Pharmaceutical companies should learn from smaller, more entrepreneurial enterprises that see value in the incremental improvements that might emerge from small-scale pilots. The experience so obtained might yield long-term benefits and accelerate the path to the future state.



Pharmaceutical companies desperately need to bolster R&D innovation and efficiency. By implementing these eight technology-enabled ways to benefit from big data, they could gradually turn the tide of declining success rates and stagnant pipelines. □

The authors wish to thank Michael Chui, Ajay Dhankhar, Sam Marwaha, Peter Pfeiffer, Navjot Singh, and Steve Van Kuiken for their contributions to this article.

**Jamie Cattell** is a principal in McKinsey's London office, **Sastry Chilukuri** is an associate principal in the New Jersey office, and **Michael Levy** is an associate principal in the Washington, DC, office.