



Big Data Requires Big Thinking

As more and more data becomes available, the role of the data scientist will rise in the industry.

Are data scientists the new superheroes of the future for pharma drug development? That depends on who you ask. The general consensus, however, is that data scientists have the capacity to improve the effectiveness and efficiency of the discovery and development process, but should work in tandem with science researchers. The optimum solution might be a melding of the work of data scientists and researchers to create a mix of traditional and real-world analysis, which would lead to more efficient development and discovery timelines.

The influence of new technologies is changing the drug development process and the assimilation of data scientists will evolve as a result.

According to a 2018 report, “Mine the Gap: How the Role of Data Scientist Fills a Need in the Pharmaceutical Industry,” authored by Michael S. Rimler, then of FMD K&L, who currently works as a clinical programming manager at GlaxoSmithKline, and Jorine Putter, director, head of programming at Grünenthal, the pharma data scientist can make positive enterprisewide contributions to small- to mid-sized pharmaceutical companies, whether directly employed, contracted

via a CRO vendor, or even as an independent contractor.

Data science can provide fast, fit-for-purpose, easily understandable knowledge and insights to questions spanning multiple data sources, by applying advanced and highly sophisticated data integration, exploitation, and analysis methodologies without losing the traceability of the data to the insight. Data science can also provide business insights based on pattern detection within vast amounts of data that is otherwise too large or too complex to visualize.

Traditionally, pharma has used data scientists who mainly consist of biostatisticians and biostatistical programmers, to ensure that sound statistical principles were applied to clinical trials. The ultimate goal of such endeavors was to gain regulatory approval of their compound based on clinical significance, substantiated by statistical significance. Over time, a need grew to apply statistical principles in the design of clinical trials to generate greater flexibility and apply adaptive designs.

Mr. Rimler and Ms. Putter say researchers then developed the need to extrapolate known effects to see what inferences could be made within different populations or subgroups.

FAST FACTS

DATA SCIENTISTS SAW A 56%

INCREASE IN JOB OPENINGS IN THE UNITED STATES OVER THE PAST YEAR.

Source: LinkedIn

This innovation led to data driven decision making, which ultimately spans almost every corner of preclinical, clinical, and commercial pharmaceutical drug development. In each of these areas, one could argue the merits for applying data science expertise, the report concludes.

Ben Wiegand, Ph.D., global head, World Without Disease Accelerator, Janssen, says he believes data scientists will definitely play a part in the success of the industry.

“Bioinformatics, machine learning, and artificial intelligence are elements of data science that will continue to play a critical role in pharmaceuticals and in the life-sciences industry,” Dr. Wiegand says. “The potential of the vast amounts of data that are generated daily, hourly, or even by the minute, have tremendous implications for how we not only manage health, but perhaps how we can maintain health and prevent, intercept, and cure disease. Considering these factors, data scientists will play an integral and paramount role in the healthcare of the future, and for the health of future generations.”

Janssen is currently using data science to help determine site selection, adds Bert Hartog, Ph.D., senior director clinical innovation, Janssen Pharmaceutica NV. And even though it holds promise in helping pharma to recruit quickly and reach underserved populations, data science is not a magic bullet, as other capabilities, such as behavior sciences, cognitive

The Typical Data Scientist 2019

Predominantly Male
(69%)



8 years
in the workforce overall



Bilingual



Python/R
(73%)



2.3 years
as Data Scientist



Master/PhD
(74%)



Data scientists will play an integral and paramount role in the healthcare of the future, and for the health of future generations.

DR. BEN WIEGAND
Janssen

engagement, and other novel technologies are still important and can be further improved, he cautions. “Data science will unlock another door to better understand disease, how disease affects people’s lifestyle and how best to design and execute clinical trials,” he adds.

According to Microsoft, data science and scientists will continue to evolve the R&D process.

For instance, the lab of the future will create new AI models and applications that will augment data scientists’ capabilities to take on the next wave of challenges in medicine, says Daniel Carchedi, senior director, business development & strategy, Microsoft.

“Data science along with AI will be the key to the future of the pharmaceutical industry,” Mr. Carchedi says. “AI will move to the desktop of every pharmaceutical employee.”

In the area of cell and gene therapy, such as individualized CAR-T therapy, the process is expensive and takes a long time and sometimes the results, while within specifications, are not consistent. Experts say by analyzing the

effectiveness of the results using AI, data scientists at pharmaceutical companies will be able to optimize the process and better predict outcomes. This will help to improve the yield and deliver consistent results each time, while driving the cost down.

“Ultimately, AI and data science could prove to be the catalyst in addressing some of today’s most difficult-to-treat health conditions,” Mr. Carchedi says. “By combining genomics with individual patient data from electronic health records and real-world evidence on patient behavior from wearables, social media, and elsewhere, healthcare providers will be able to harness the power of precision medicine to determine the most effective approaches for specific patients.”



AI and data science could prove to be the catalyst in addressing some of today’s most difficult-to-treat health conditions.

DANIEL CARCHEDI
Microsoft



Data science will unlock another door to better understand disease, how disease affects people’s lifestyle, and how best to design and execute clinical trials.

DR. BERT HARTOG
Janssen Pharmaceutica

Data Scientists: Tomorrow’s Pharma Superhero?

Experts express their views on the growing role of data scientists in the industry.



SARAH ALWARDT, PH.D.
VP, Operations
McKesson Data,
Evidence & Insights,
McKesson

A few facts on data: there are 2.5 quintillion bytes of data created every day. The amount of data we are creating is doubling every two years. There are estimates that suggest that only 0.5% of these data are ever analyzed. Taking a pragmatic look at the data sources we have today, the data essential to drive innovation, and speculate about the data we don’t even know we have or need yet, truly boggles the mind. If we believed that EMR or claims data were complicated, when we add in terabytes of genomic, proteomic, and exomic data and then add in wearable data, social media, unstructured data, and data on every interaction we have, we have more than our brains can hope to process. Data scientists can unlock hidden patterns in these vast data pools. They will be the ones who will create predictive models that can tell individuals whether a new drug will work for them. Mining data will enable us to better understand clinical outcomes and

put together large and disparate sources to allow us to draw meaningful conclusions and insights. If we extrapolate to R&D and think about finding the next new cell and gene therapies or CRISPR therapies, only data scientists will be able to discover genetic drivers of mutations that cause cancer while sifting out the genetic passengers. Their abilities with data will underscore new development in these exciting therapeutic categories.



MARCO ANELLI, M.D.
Principal Consultants
Coordinator and Head
of Data, Information,
Knowledge & Intelligence
Group, ProductLife Group

Across all stages of the life cycle, pharmaceutical companies have massive amounts of data. Thanks to advances in simulation and other technology, it has become cheap and easy to run an experiment, which results in a lot of discovery data. In clinical trials, new technologies such as smartwatches mean far more data is generated, a lot of which could be useful to determine the efficacy of a drug. In pharmacovigilance, vast amounts of data are being generated that can help companies to find

out if patients have had relevant adverse events. Insurance companies are collecting every piece of information about a patient’s health for good pharmacoeconomic reasons. The result is pharmaceutical companies have massive amounts of data, much of which could be extremely valuable in gaining insights, and they must find ways to reduce the data they are working with to gain those relevant insights. This is known as dimensionality reduction. This is precisely why data scientists are so crucial. Indeed, perhaps one of the best descriptions of data science is “data that is too big to be handled in the traditional way.” Data science reduces data to something that humans can use and make sense of, while ensuring that relevant information is retained. The industry will need to exploit data science to be more efficient and effective and to improve the bottom line.



HELENE BRASHEAR
Senior Data Scientist,
Data Sciences, W20
Pharmaceutical compa-
nies generate diverse data
beyond clinical data. Many
functions within a com- ▶

Data Scientists: Tomorrow's Pharma Superhero?

▶ Many generate data — ranging from scientific discovery to manufacturing logistics to marketing. While all of these functions are connected within an organization, the data they generate is often siloed. Data scientists are critical for the industry because they help organize data and connect both business needs and insights to that data. Companies that leverage data science teams are more informed, agile, and better able to make holistic decisions for the entire organization.

Successful data science initiatives use interdisciplinary teams that bring together skills from engineering, programming, and data modeling to manage the data lifecycle. Clean, organized data lays the foundation for analytics. Analytics and data modeling generate understanding of the data contents, trends, and behaviors. These insights provide measurable feedback for business processes that enable informed decision-making. Siloed data requires a team that can wrangle the data as well as understand what the data contains. These interdisciplinary teams are communicators that collaborate with experts in silos to share information and information context needed across the organization.

Additionally, data science is the foundation required to build technology systems for AI and machine learning. Machine learning requires clean data, data pipelines, and an analytical understanding of research questions. Quantitative models and scientific data provide opportunities for innovation using classification, inference and prediction. Business process data helps identify lucrative opportunities for machine learning in areas such as manufacturing, marketing, and supply chain management. Good data science is the glue that connects key functions within a company and enables the organization to effectively use the data it has while planning for the future.



JOE DEPINTO
President, Specialty Solutions, Cardinal Health

Data scientists will play a critical role in the future of healthcare. The pharmaceu-

tical industry is increasingly relying on data to drive decision-making in clinical drug development, commercial launch preparation, patient adherence, and value-based care contracting, just to name a few areas. The recent advances in AI have enabled the industry to monitor, query, and manage big data in ways that have been difficult and time consuming in the past. As AI tools become more advanced, the predictive value of this work has the potential to change the outcomes of care. However, the tools cannot stand alone — data scientists will be needed to review, analyze, and challenge the findings to ensure they address the intended goals. In the past, much of this data science work has been managed in academic settings. Today, with the increased demand for data analysis to support clinical development and business strategy, pharma and biotech companies are building data science capabilities internally, or outsourcing the work to partners, or both. The most effective data science teams will be those that not only understand the tools, but also understand the clinical implications, the value/cost benefit, and how the data can be applied to optimize patient outcomes.



BRENDAN GALLAGHER
Chief Connected Health Officer, Digitas Health & Publicis Health

Our ability to meaningfully gather, organize, structure, and use data will be the key to almost every industry but maybe most particularly the pharmaceutical industry. Vas Narasimham, CEO of Novartis, famously declared in 2018 that they would lead the revolution, positioning Novartis as a medicines and data company adding, "We're using machine learning and predictive analytics to improve how we initiate and manage clinical trials making trial timelines shorter, while improving the quality of data we collect." This could have a massive impact on the cost of bringing new drugs to market. Combine that with the efficiencies data science brings to the operational side of delivering those medicines, such as eliminating waste and personalized care, and we have the makings of a transformed healthcare marketplace. The pharmaceutical industry is one of the

only industries where data science has the potential to disrupt both the R&D and the commercial side of the industry in profound ways.



CHRIS GARABEDIAN
Chairman & CEO, Xontogeny

While I agree with the sentiment of medicinal chemist and industry blogger, Derek Lowe, who stated "Machines won't replace chemists, but chemists who use machines will replace those who don't," I believe the machines will still have a less important role than the chemist, in this example, or an experienced biopharma professional. While the most fruitful area of pharmaceutical development in which data scientists and AI-driven algorithms have manifested is in the areas of drug discovery and viable lead screening, the value in advanced drug development will be far more elusive. Just as high-throughput screening did not result in a larger yield of drugs that outperformed in clinical trials, it is unlikely that screening of DEL-libraries with trillions of compounds will solve the 95%+ failure rates between IND candidates and FDA approval. It is true that AI and data scientists may drive better starting points with molecules in which we understand drug-like characteristics, it will unlikely materially change the success probabilities of clinical trials. There are far too many variables, many that have little or no precedent, which influence clinical trial success. First, the database of all clinical trials contains mostly failed studies which can be hard to deconstruct and identify the driver(s) of failure. The universe of successful clinical trials is far smaller, especially when selecting for specific therapeutic areas or disease indication. Second, each drug is unique with a different mechanism, modality, dose, activity profile and is often studied in a unique sub-population of a broader disease. Lastly, clinical trial design choices have endless permutations and are not easily weighted or quantified and require experience-based subjective judgments around the best application in patients, clinical endpoints and biomarker assessments, ▶

Data Scientists: Tomorrow's Pharma Superhero?

► selection of dose/regimen/formulation, inclusion/exclusion criteria, powering assumptions and often with regulatory constraints and guidance as a constraining factor.



MICHELLE HOISETH
Chief Data Officer,
Parexel

Data scientists will absolutely be key to our future in the development and delivery of new therapies. Not just any data scientist, though. They need to be trained in the data that matter to us in the life sciences. The other gap that is becoming increasingly apparent is that knowledge of the data and options to prepare data for analysis fall short in the absence of domain experience in the area of clinical development in which those data are intended for use. We as an industry are facing a substantial skills shortage in this area. It is important that we support an increase in the formal training of these people, and equally critical, start to have a serious discussion about how we are going to make skills acquisition and professional development available to those in our organizations who need and want to acquire them. Failure to do so will end up being the thing that delays our ability to create the change so important to long-term success in clinical development. Overall, data scientists are a critical component to the success of our industry as we all work to deliver important new medicines to patients in need.



DR. LAN HUANG
Co-founder, Chairman,
CEO, BeyondSpring

It's been more than 20 years since we have sequenced the human genome and the industry has successfully incorporated biomarkers into selection for both drug development and clinical use. Structural biologists, in addition, can share with us why certain drugs work and other don't from a molecular perspective. The hope is that we will be able to efficiently categorize response and safety data

to differentiate classes of drugs for specific diseases. Data scientists will build huge, clinically focused databases, which will not only assist in developing new drugs, but also in identifying the best patients in the clinic. The end result: reduced costs, improved care for patients, and better outcomes.

ALAN KOTT
Practice Leader Data Analytics,
Signant Health

While I believe that data scientists will continue to play an ever-increasing role within clinical trials, I do not believe that they'll be the primary role, or "key." In terms of clinical trials, data scientists are always great in enhancing information, finding new ways to translate or interpret data, and uncovering new angles. In fact, they are responsible for several of the industry's new developments and areas of interests.

However, with all that being said, data scientists will always lag behind clinicians when it comes to clinical trials. In my belief, clinicians will always play the primary role within trials, with all other roles aiming for second place. Imagined another way, clinicians will always be the primary tool for trials, with data scientists finding new ways to maintain and use the tool.



AMIR MALKA
President and
Co-Founder,
BioForum

Pharma's focus on data science isn't new, if you consider biostatisticians and biostatistical programmers who have long ensured that sound statistical principles are applied to clinical trials. Today, however, we're seeing an explosion in the amount of available information, beyond what's collected in traditional, tightly controlled clinical trial settings. Advances in technology and data science techniques are paving the way forward for the industry, enabling us to unlock insights that we can use to get better drugs to patients faster.

While it was once sufficient to analyze relatively small amounts of clinical trial data, pharma now needs skilled people and partners, along with cutting-edge platforms, to help extract insights from

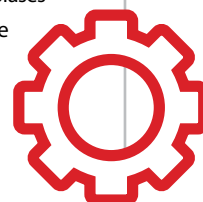
vast amounts of available real-world data, including unstructured data such as physicians' notes, images and lab reports. These data sources come in a variety of different formats and are messy, filled with inconsistencies, potential biases and noise. While much still needs to be done to create standardized methods for sharing and making sense of real-world data, it's now possible to link different data sources, which provides an opportunity to better understand diseases, treatment patterns and clinical outcomes in an uncontrolled, real-world setting. It's also now possible to learn more about a wider spectrum of patients, including those who are traditionally excluded from clinical trials.

Consequently, across pharma, there's not only a tremendous need but also a huge opportunity for data scientists to make an impact. Those who can figure out how to effectively apply modern data science techniques and tools to drug development, in a thoughtful and fit-for-purpose manner will play a key role in transforming clinical research and bringing needed new therapies to patients all over the world.



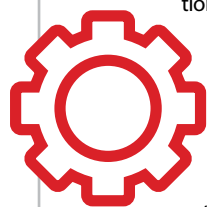
THOMAS NEYRAPALLY
Chief Commercial
Officer, Sema4

The work of data scientists will enable better decision-making across the pharma value chain given the inherent limitations in the human brain's ability to take into account more than a small number of variables and data points at the same time in making decisions. Data scientists will also enable better connectivity between parts of pharma companies that have not traditionally worked in as well integrated a manner. For example, patient journeys have become of great interest to everyone from early drug discoverers to marketers, and data scientists can work to serve up the relevant data and the relevant views to the disparate constituencies within pharma interested in these journeys. Data scientists can also enable better connectivity between pharma ►



Data Scientists: Tomorrow's Pharma Superhero?

► and its constituencies — patients, providers, and payers — by providing information to support decision-making. During the development of a drug, large quantities of data are generated, yet traditionally only a fraction of that data exhaust ends up being utilized. Data scientists can drive the extraction of significant value from this data. While such data has often been left on the shelf in clinical development, data scientists may be able to use that data to build computational models of disease to inform new drug targets and effective drug combinations. Despite all of this potential, there remain some critical limitations in the ability of data scientists to drive the future in pharma. Data scientists do not possess magical powers if the appropriate data is not collected to address a given question, and even the most powerful data science tools will not be able to conjure up an accurate answer. Critically, data scientists will need to become more integrated with the functional teams within which they work, rather than remaining in separate data science groups, and will need to undergo significant cross-training in the functional areas with which they work.



KATHRYN TICKNOR
Head of Research,
inVibe Labs
@invibelabs

In pharma, our driving focus is how people come together to make decisions about their lives and bodies. These represent some of our most important, most intimate, most life-changing moments: as a result, social scientists, not data scientists, will be key to the future of our industry. Long maligned with words like “softer” or “subjective,” social science is quickly gaining traction as a critical element within pharma’s ongoing research priorities. This recent buzz is well-deserved: social scientists can unlock a wealth of critical insights to help inform pharma’s understanding of how people navigate the healthcare ecosystem. Don’t be fooled — social scientists are

no strangers to data. Indeed, they are experts at analyzing some of the most complex, “messy” datasets there are: patient voices; natural language; human emotions. Like biologists peering into a microscope, social and behavioral scientists have developed a professional vision for spotting patterns through the lens of human thought, language, and behavior that would otherwise be missed entirely. From optimizing marketing strategies, to rethinking clinical trial development, and redesigning patient support services, pharma knows that initiatives must reflect a more fully contextualized, personalized understanding of their customer’s identity and unmet needs. To remain relevant among the shifting tide of consumer-driven healthcare, the expertise of linguists, anthropologists, and social scientists will be key.



KEVIN TROYANOS
Head of Analytics,
Publicis Health
@PublicisHealth

In a 2015 study, Forbes Insights found that 60% of executives were “expect[ing] to increase their spending on data-driven analytics over the [following] three years.” That same year, Gartner released a seemingly contradictory study claiming that, “through 2017, 60% of big data projects will fail to go beyond piloting and experimentation, and will be abandoned.”

Here we are nearly a half decade into the future, and the industry leaders who doubled down on data science — in keeping with Forbes’ predictions and seemingly at great risk — now appear quite prescient. In today’s pharma landscape, data science not only shapes the work of formally trained business analysts, but the work of a wide range of professionals across R&D, commercial operations, patient services, and perhaps most of all, marketing.

And yet, many executives still struggle to effectively translate data science-driven insights into action, largely because data science is often viewed as work wholly owned by data scientists. As such, baseline data science literacy has become an all but essential trait of the modern pharmaceutical executive in deploying activation-ready data science solutions. As with any language, fluency in data

science requires time and immersion. The important thing for pharmaceutical business leaders to remember is that beyond the intimidating walls of code, models, and algorithms — data science is only valuable when tied to the “real world” in some way, shape, or form.

So, will data scientists be the key of the future of the pharmaceutical industry? Maybe. But it will be the business leaders who are able to learn the language of data science and translate the algorithms into action that will truly change the way the industry operates.



YILIAN YUAN
Senior VP, Data Science
and Advanced Analytics,
IQVIA

The role of data analytics is changing the pharma industry. With large and growing data sets — from R&D, patients and retailers — there is an increased demand for data scientists to analyze and draw insights from this information and to be able to do it faster. We know bringing drugs to market is a time-consuming and expensive process. By leveraging data modeling and predictive analytics, clinical development processes can become more efficient and cost-effective. These capabilities help researchers and the industry make more informed decisions.

Data science is a demanding discipline that requires the ability to analyze large amounts of data and to identify patterns and deep insights. For the pharma industry, data scientists play a pivotal role in being able to help improve healthcare. Through their expertise, they can help partners identify drugs that show the greatest market promise, identify the patients who might best benefit from their products, and deliver personalized care to patients, analyze real-world data to track and avoid adverse events, and mitigate potential risks to patients.

As the data revolution evolves, data scientists will be critical in helping the industry realize efficiencies that will substantially transform the way clinical research is done, and pharma companies and, most importantly, patients will be the beneficiaries. **PV**

(c) PharmaLinX LLC. Rights do not include promotional use. For distribution or printing rights, contact mwals@pharmavoices.com

Be
heard.

Be here.

Pharma **VOICE**

THE FORUM FOR THE INDUSTRY EXECUTIVE