



## *Bringing Analytics and AI into the Clinical Setting*

Like any hospital, the Medical University of South Carolina (MUSC) wanted to reduce the number of patients developing sepsis. To do that, MUSC first needed a better way to find those most at risk of developing the life-threatening condition in response to an infection.

Using data analytics and machine learning (ML) modeling, MUSC was able to capture livestreaming data from electronic health records (EHRs) and real-time telemetry instruments in the intensive care unit. These records were used to build better classifiers for [identifying patient deterioration](#) due to sepsis. Later, MUSC ported that same automated data analysis program to get a better read on COVID-19 patients, specifically those who were at greater risk of decline.

The South Carolina health system is among the healthcare providers everywhere realizing the potential in harnessing the petabytes' worth of digital information routinely generated within today's clinical settings. With the help of unified data analytics and artificial intelligence (AI) platforms, healthcare providers are streamlining workflows, enhancing precision health, expanding population health and providing more patient-centric care.

"There's a great opportunity to use this data to better understand how patients are engaging with their care and to make alterations and suggestions that can lead to better utilization of their benefits and, ultimately, improved healthcare outcomes," explained Michael Ortega, Industry and Solutions Marketing Leader at Databricks, a global leader in data and AI.

The COVID-19 pandemic has forced many healthcare organizations to rethink their IT strategy so they can marshal an agile response to the changing conditions of the pandemic. These IT strategy changes can simultaneously drive improvements to the clinician and patient experience, patient outcomes and the financial futures of the healthcare organization.

## Opportunities to elevate healthcare

The opportunity for data and AI to improve patient outcomes and reduce the cost of care is monumental. There are three major areas where analytics in combination with AI and ML are transforming healthcare.

**Streamlining administration:** Healthcare organizations can often systematically examine their data to streamline the administrative side of care, particularly to optimize patient flow and workflows within the clinical setting. These operational efficiencies leave more time for direct patient care and provide financial incentives. For instance, health insurers now use analytics and AI to reduce fraud, waste and abuse. Similar techniques can be used in a hospital to automate revenue cycle management.

**Unlocking precision care:** Health systems are also applying these tools to improve individual care plans, pulling data from myriad sources like EHRs, genome sequencers and medical imaging devices for earlier diagnoses and targeted treatments. In one case, the precision medicine company Human Longevity combined sequencing and whole-body MRI data sets to build a classifier that used specific [biomarkers to identify dementia](#) symptoms in patients years ahead of more standard diagnostics. Other examples include CVS Health,

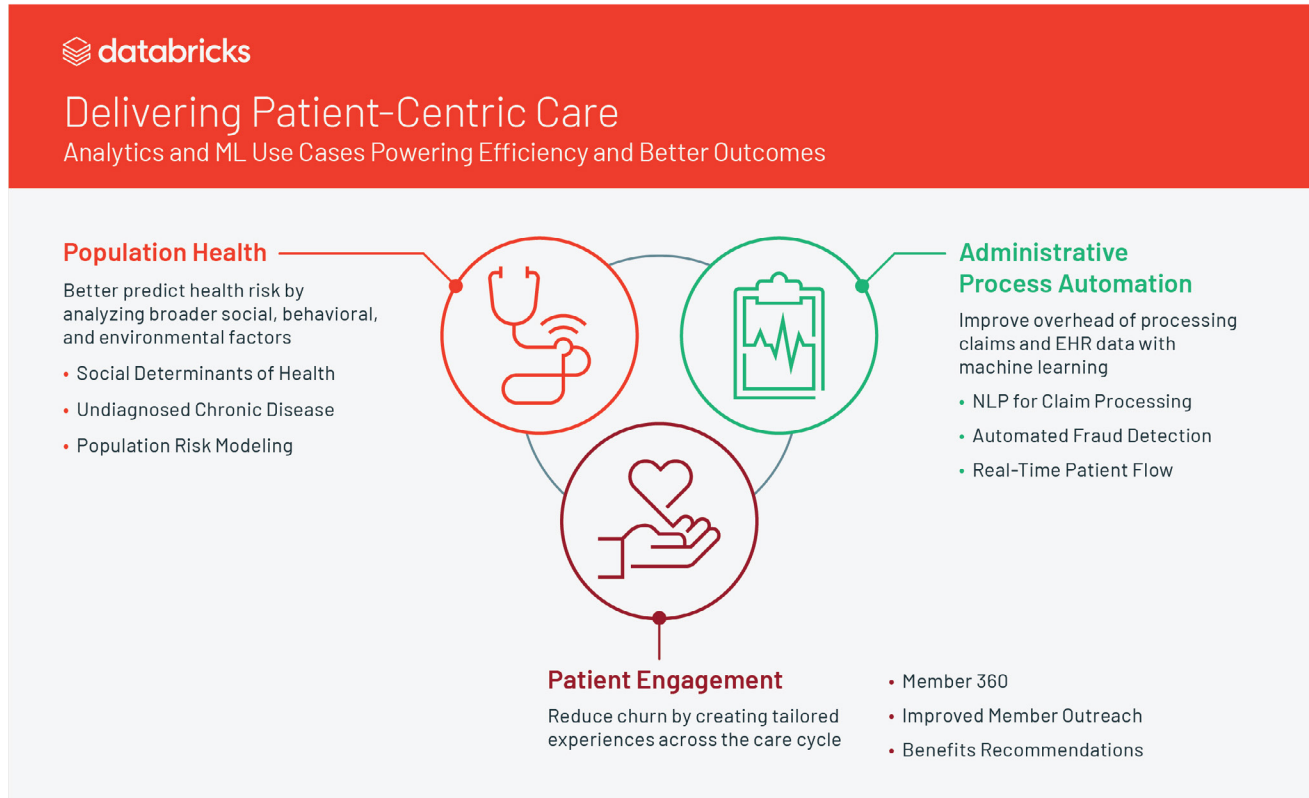
MUSC and Regeneron. Each of these companies used AI, machine learning and/or predictive analytics to improve outcomes for millions.

**Improving patient engagement:** Digital technologies have greatly improved how patients communicate with their providers. Chatbots, personalized alerts through wearables and online portals with recommended content are a few areas where AI is enabling a more personalized patient experience. As an example, CVS, a nationwide pharmacy chain, [examined 70 million prescription](#) records to help decide which patients were at higher risk of contracting COVID-19 and, therefore, should have medications mailed to them rather than pick them up in person. This improved treatment adherence by 1.6%.

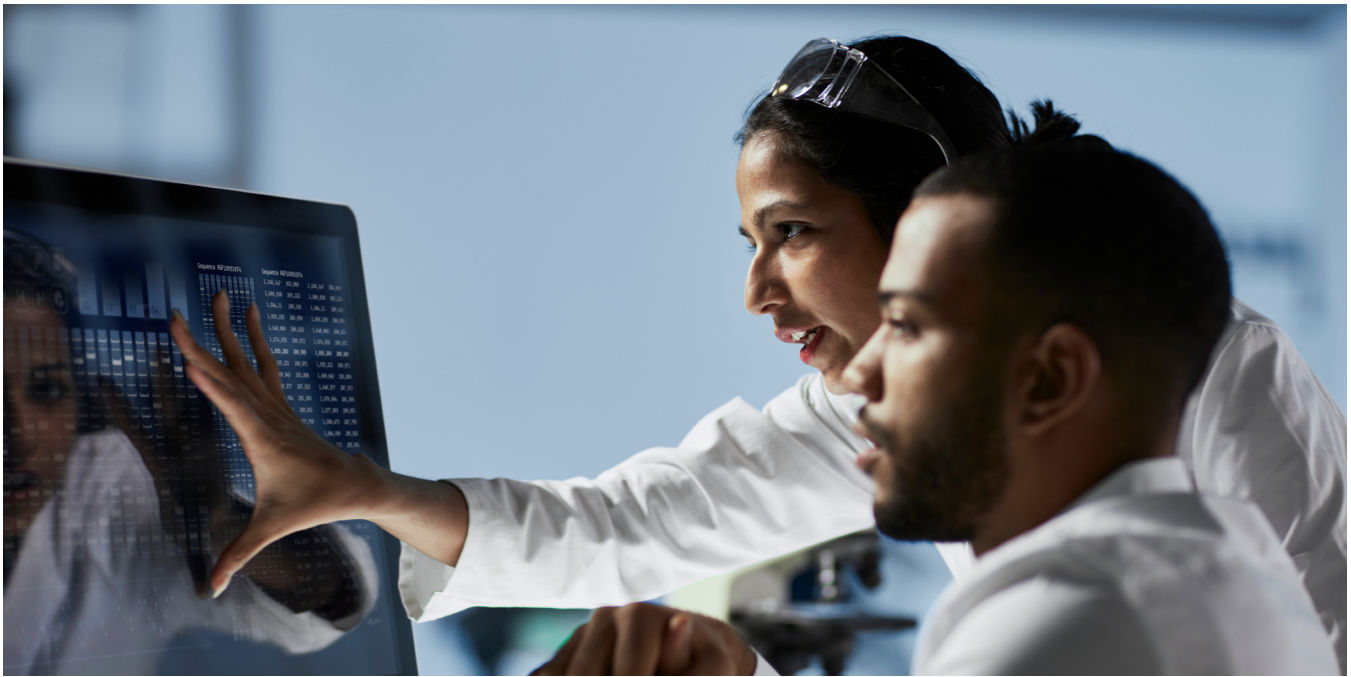
In all these instances, organizations relied on a platform with tremendous processing power and storage capabilities that can unify health data from a wide range of sources into models that generate impactful results.

“Ultimately, by blending all these data sets together and coming up with an end-to-end way to engage with patients, both in the hospital and outside, we can ultimately improve their health outcomes and drive better usage of our healthcare services,” said Bavesh Patel, Vice President of Industries at Databricks.

**Figure 1.** Some of the ways data analytics is used to improve efficiencies and outcomes







## Common barriers along the data analytics journey

Every organization has unique challenges. But they also encounter common obstacles when deploying data analytics programs, such as the quality of patient data gathered for ML opportunities. Other issues routinely arise from:

- Siloed data that's difficult to integrate
- Scalability due to the volume and velocity of data sets
- Model reproducibility and making models accessible in a clinical setting

Databricks provides a Unified Data Analytics Platform with robust storage and computational capabilities in the cloud. The platform makes use of popular open-source software like Apache Spark™ and the Delta Lake storage format, thereby funneling large amounts of healthcare data into a single lake upon which to conduct analyses. Information from EHRs, genomics, imaging and many other data sources is loaded more quickly using data interoperability standards like [HL7](#) and [FHIR](#) (Fast Healthcare Interoperability Resources) and tools like [Glow](#) for genomic data and [DICOM for radiology](#).

“Ultimately, with this data all stored in a single environment, we can blend [everything] together very rapidly with a very low cost of storing the data,” Patel explained.

Additionally, to improve replication of modeling within a clinical setting, there's [MLflow](#), an open-source project founded by Databricks that is natively integrated into the

platform. MLflow provides an end-to-end system for logging, storing and validating ML artifacts. Models built by one data science team are logged and stored in a model registry that can be accessed by other teams working on similar projects. This improves confidence in newer models while reducing the time it takes to test and deploy them.

Furthermore, pre-built templates based on best practices and health data regulations, like HIPAA/HITRUST, make it easier for healthcare analytics teams to roll out their own use cases with build times reduced from months to weeks in some cases. At Sanford Health, for example, these capabilities have helped reduce the time it takes to turn clinical research into a validated screen ready for use in the clinic.

The goal of all these offerings is to develop high-value ML opportunities in areas like population health, administration and patient engagement that deliver lower-cost patient care with better health outcomes, while eliminating IT barriers through a streamlined cloud platform that provides strong security, and state-of-the-art data and model governance.

***The goal of all these offerings is to develop high-value ML opportunities in areas like population health, administration and patient engagement that deliver lower-cost patient care with better health outcomes, while eliminating IT barriers through a streamlined cloud platform that provides strong security, and state-of-the-art data and model governance.***

## UCLA Health: Building a strategy for modern cloud analytics

UCLA Health in Southern California is among the healthcare systems adopting modern cloud analytics to provide such patient-centric care. It now routinely pulls myriad health data into one comprehensive patient medical record that's easily and securely retrieved in the course of care. Clinicians do not waste time tracking down data housed in different systems, and patients can use patient portals to access lab results and communicate with their care teams.

That investment in a more patient health-centered data program was done primarily to foster greater clinician and patient interaction. But it also paid off once the coronavirus struck and priorities and patient loads shifted. For example, telehealth and telemedicine became more widely available – and popular. The hospital went from seeing about 400 patients monthly via video sessions to more than 60,000, said Michael Pfeffer, MD, UCLA Health's Chief Information Officer and Assistant Vice Chancellor as well as an internal medicine physician.

UCLA Health's strategy for building a modern cloud analytics program was predicated on four key areas:

- Aligning clinicians and data scientists around shared objectives
- Enabling use cases through data unification
- Identifying high-value ML opportunities
- Enabling best-of-breed technologies

"It's really helping us as a health system take better care of patients on a more personal level, on a more individual level, whether that's from how they access our systems to actual

clinical decision support," Pfeffer said during a recent HIMSS webinar titled "AI and Healthcare: How to Bring Analytics and AI Into the Clinical Setting."

When the pandemic hit, a computational medicine team from UCLA's medical and engineering schools had a modern analytics program in place. This allowed the team to rapidly build models to determine where COVID-19 infections were most likely to occur. Then, the hospital could quickly adjust hospital operations – such as scaling down certain surgeries – to meet shifting demands with COVID-19 patients.

UCLA was among the first hospitals to develop successful COVID-19 testing and make aggregate results publicly available on its website. The hospital applied ongoing data analytics to reallocate resources based on increasing or shrinking caseloads. Now it's considering adding an AI layer to help predict likely COVID-19 outbreaks and outcomes in different populations it serves.

One ongoing issue for the UCLA computational team is incoming inaccurate or incomplete data in need of repair or cleansing before it can be used. This is especially difficult if the problem is subjective documentation. Consider how "length of stay" can be widely interpreted due to discrepancies in noted discharges. Is a discharge duly recorded when the order is written? When a timestamp is given? Or when the patient leaves the hospital?

That's where good data governance comes into play.

"Our research community has been incredibly helpful in identifying where there are opportunities and quality control on the data," Pfeffer said. "I think that partnership is really, really critical. Our physician informaticists and our nurse informaticists kind of stand in the middle of helping the business understand our data scientists, understand where those bridges are, and can then review workflows to determine where we can do better or help interpret the data, which is actually very, very complicated."



## Lessons learned from big data experiences

Based on his organization's experiences, Pfeffer offered the following tips to those either launching a data analytics program or currently stuck in the weeds.

*Have the right (and tight) collaboration in place.* Carefully choose key stakeholders and trusted partners to create a vision and sound strategy around it. Then provide the appropriate level of IT oversight and data governance to make sure next steps are understood and universally followed. Working collaboratively, teams are more apt to stay on track and discover new opportunities.

*Secure a talented data analytics team.* A data analytics team should be comprised of people passionate about data science, analytics, visualizations and the data quality within the IT organization. It's their talent and dedication that build momentum, overcome obstacles and yield results.

*Focus on implementation.* Results of modeling or any analyses need to be implemented by physicians, clinicians, nurses and others who provide direct patient care. Make sure you can explain not just what results reveal, but what's in it for the practitioner. That's how you'll get the broad buy-in now needed. Make sure you present information in a user-friendly way, too.

*Beware of algorithm biases.* People may not know how to create sophisticated algorithms, but they increasingly are aware of inherent flaws. Bias can infiltrate algorithms in different ways. The most important thing is to remain aware of programmers' predispositions so any biases can be mitigated before the algorithm is put into production.

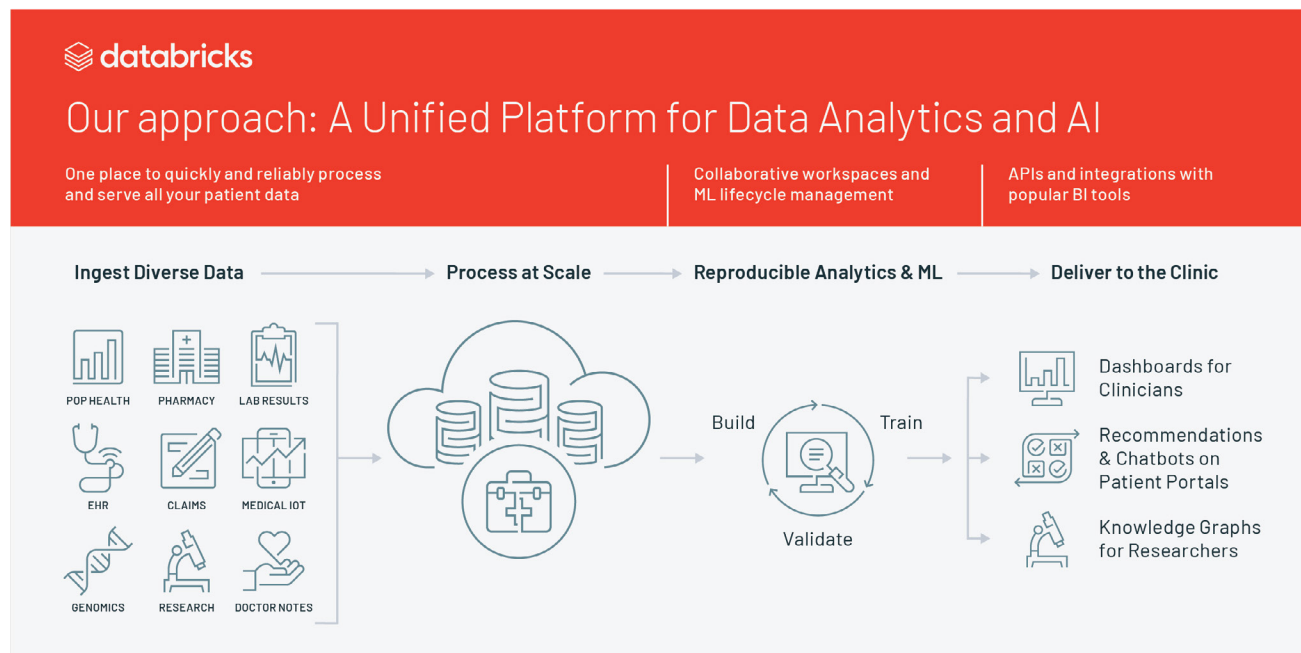
*Validate. Validate. Validate.* Sharing data models is critical, not just for learning but also for reproducing something similar in a clinical setting. Understand that all models must be tested, then tweaked and tested again. Rather than assume you can plug and play, think: plug, test, play, plug, test, tweak, etc.

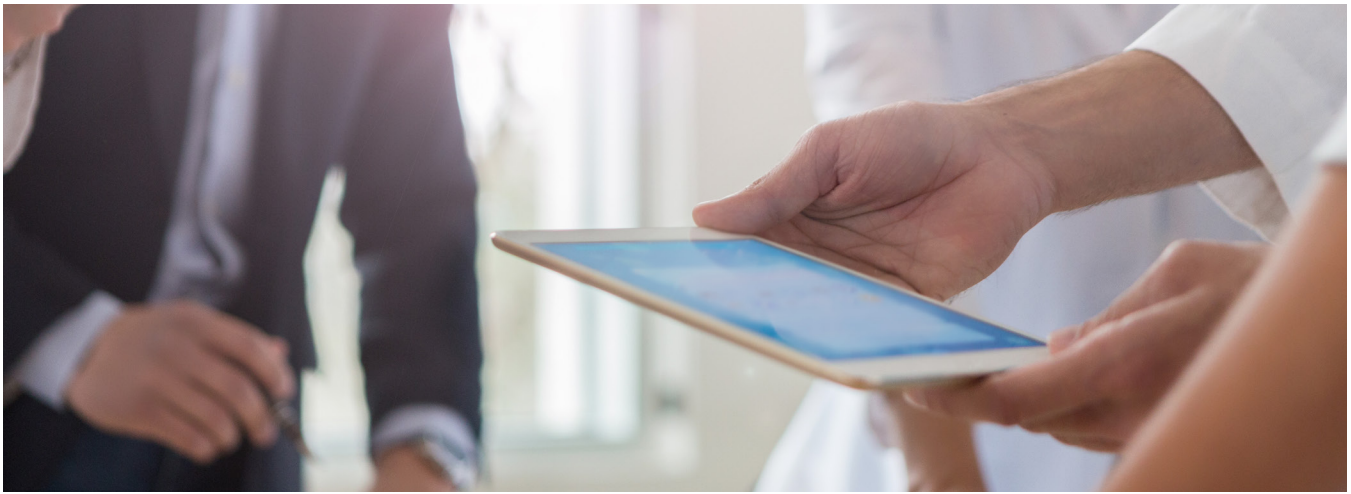
## The path forward to patient-centric care

An open, unified data analytics platform – like the one offered by Databricks – that considers the entire data and machine learning lifecycle can be a catalyst for revolutionizing patient care (Figure 2). These technologies remove some of the common barriers to developing ambitious big data programs, given their ability to ingest diverse data and process it at scale. They also provide replicable modeling for future ML use cases. And, they help informatics and data science teams generate reports and easily understood dashboards for clinicians charged with implementing whatever new processes or procedures come from those findings.

***Results of modeling or any analyses need to be implemented by physicians, clinicians, nurses and others who provide direct patient care. Make sure you can explain not just what results reveal, but what's in it for the practitioner.***

Figure 2. The Databricks approach





Additionally, advanced AI tools can improve patient engagement, such as through chatbots on well-designed mobile patient portals, or add to clinical research tied to groundbreaking treatments or cures. When applied to clinical settings, these programs can fulfill an unmet clinical need or align resources to fill in gaps in care.

“At Databricks, one of our four key cultural values is to let the data decide,” Patel said. Focus on what’s necessary from an end-to-end system design and, ultimately, how results of queries are presented to clinicians.

Pfeffer recommends always staying focused on the patient as a guiding principle.

“Everything we do is about the patient,” he said. “We want to make sure that we’re actually solving problems that will improve patient care, reduce health disparities and improve our ability as an organization to take care of people, which is operations.”

Not every problem can be solved with AI and ML, he added. It may be something as simple as tweaking a workflow or providing more education. Also, an algorithm may perform only slightly better than a traditional scoring system.

“Doing this stuff is not easy. The platforms are complex, and data security is incredibly important,” Pfeffer said. “With all that in mind, we really want to focus our energy and resources on solving those really important problems that benefit our patients because that’s why we’re all here. I mean, that’s what makes this all so fun and amazing, that the problems we’re trying to solve directly impact patient care.”

---

**To learn more about the power of data analytics and artificial intelligence in healthcare, visit [databricks.com/healthcare](https://databricks.com/healthcare).**



#### About Databricks

Databricks is the data and AI company. Thousands of organizations worldwide — including Comcast, Condé Nast, Nationwide and H&M — rely on Databricks’ open and unified platform for data engineering, machine learning and analytics. Databricks is venture-backed and headquartered in San Francisco, with offices around the globe. Founded by the original creators of Apache Spark™, Delta Lake and MLflow, Databricks is on a mission to help data teams solve the world’s toughest problems. To learn more, follow Databricks on Twitter, LinkedIn and Facebook.