Health & Medicine

How Artificial Intelligence Is Changing Biomedical Research

Columbia researchers are using AI to unravel genetic mysteries behind cancer, Alzheimer’s, and more.

By
David J. Craig
|
Spring/Summer 2025



Photo illustration by *Columbia Magazine*. (Adobe Stock)

Gone are the days when biomedical researchers spent all their time scrutinizing test tubes, culturing microorganisms, and counting cells by hand. Today, they’re as likely to be poring over computer spreadsheets as peering into microscopes.

Why? Consider this: We now know that the human body is made up of thirty-seven trillion cells, all with distinct roles, patterns of genetic activity, and physical characteristics. Inside every cell, roughly twenty thousand genes and up to ten thousand proteins collectively engage in millions of interactions per second. And no cell operates in isolation; each is constantly chatting with its neighbors and adapting to its environment.

“To make sense of this complexity requires collecting large amounts of data and identifying patterns within it,” says Raúl Rabadán, a Columbia professor of systems biology and director of the department’s program for mathematical genomics. “Biology is increasingly becoming a quantitative field.”

A key driver of this shift is the rapid evolution of artificial intelligence, which is allowing scientists to leverage a flood of biological data that has been generated since the early 2000s. As AI and technologies like genetic and proteomic sequencing continue to advance, biomedical research is entering a new era — one where the body can be viewed as a vast, interconnected system of data waiting to be interpreted.

At Columbia, large multidisciplinary teams of biologists, physicians, computer scientists, and mathematicians are now collaborating on data-driven projects that could transform our understanding of human health and disease. Some are working to improve treatments for diseases like cancer or Alzheimer’s; others are building AI-powered tools that could enable researchers without expertise in data science to gain deeper insight into everything from prenatal development and aging to diet and nutrition.

The shape of things to come

One of biology’s greatest mysteries is how cells — despite carrying identical DNA — take on different roles in the body and adapt to changes in their surroundings over time. For example, immune cells alter their shape to engulf pathogens, while other cells can adjust their rate of duplication based on environmental cues. Scientists have long understood that this flexibility comes from cells’ ability to switch genes on and off, thereby altering protein production and shaping their identity and function. Yet the internal mechanisms that drive these epigenetic shifts — and the factors

that determine whether they proceed smoothly or go awry — remain elusive due to their overwhelming complexity.

Rabadán's team recently developed a technology that could help illuminate these processes. The researchers, by training computers to sift through data from millions of human cells, created an AI program that can predict how genes, proteins, and other molecules in any given cell are likely to interact, based on how they've been observed to interact in other cells in the past. "In the same way that ChatGPT learns how words should fit together in a sentence, our model learns how cellular components tend to behave and respond to one another," says Rabadán.

Using the new tool, Rabadán says, scientists can dramatically increase the speed and efficiency with which they study molecular networks. Called GET (short for General Expression Informer), the open-source model enables scientists to test out large numbers of hypotheses in silico before committing to time-intensive lab experiments. "You can use the model to simulate lots of molecular interactions and identify the most promising possibilities to study with traditional methods," he says. One potential application is for investigating how cells regulate gene expression. "It provides a powerful new method for studying the most fundamental questions in epigenetics," says Rabadán. "Like, how do stem cells transform into specialized cells? How do immune cells know when it's time to attack? How do healthy cells turn cancerous?"

While other research groups have created AI tools to model activity within specific types of cells before, the new Columbia tool is the first to identify overarching patterns of molecular activity across all major cell types. Xi Fu, a PhD student in Rabadán's lab who led the creation of GET and is now working to improve its predictions, says the team's ultimate goal is to uncover universal principles that govern cellular behavior — something akin to "Newton's laws of biology."

The potential medical applications of GET are already coming into focus. In a January 25 paper in the journal *Nature*, Rabadán, Xi, and several colleagues announced that they had used the tool to expose previously unknown regulatory mechanisms behind an inherited form of pediatric leukemia. "These kids inherit genes that are mutated, but until now nobody knew what they did," says Rabadán, who also co-leads the cancer genomics and epigenomics research program at Columbia's Herbert Irving Comprehensive Cancer Center. His team mapped how the mutations disrupt normal protein interactions and contribute to the disease, a discovery that could help

scientists in their search for new treatments.

In recent months, several other Columbia medical research teams have unveiled new computer-based tools, although most still need refinement before they can be broadly applied in clinical research. Computational biologist Elham Azizi and her team, for example, have developed a machine-learning program that describes how immune cells and cancer cells adapt to each other in their struggle for survival. The researchers say the tool, called DIIsco (for Dynamic Intercellular Interaction scRNA-seq), could eventually be used both to advance research into the human immune system's capacity for fighting cancer and to guide treatment strategies for individual patients. "Right now, we don't have any reliable methods of tracking how cells evolve in response to each other over time, but that will be essential for developing better immunotherapies," says Cameron Young Park, a PhD student in biomedical engineering who is helping to lead the project.

Meanwhile, a team led by Columbia biostatistician Zhonghua Liu has built an AI model to identify the key genetic drivers of diseases with complex genetic profiles. In a recent paper in *Cell Genomics*, he and his colleagues reported that they have used the tool to pinpoint seven genetic mutations that may contribute to Alzheimer's disease more significantly than previously known. Notably, several of the mutations appear to trigger molecular damage that could potentially be mitigated by drugs already approved by the FDA for other conditions.

According to Rabadán, these breakthroughs are only the beginning. As AI grows more sophisticated, he says, the next major milestone will be predicting cellular changes before they occur — a shift that could propel personalized medicine further into the realm of prevention.

"We are entering an exciting new era," he says. "Biology is being transformed into a predictive science."

Read more from ***David J. Craig***



Guide to school abbreviations

All categories >
Read more from ***David J. Craig***