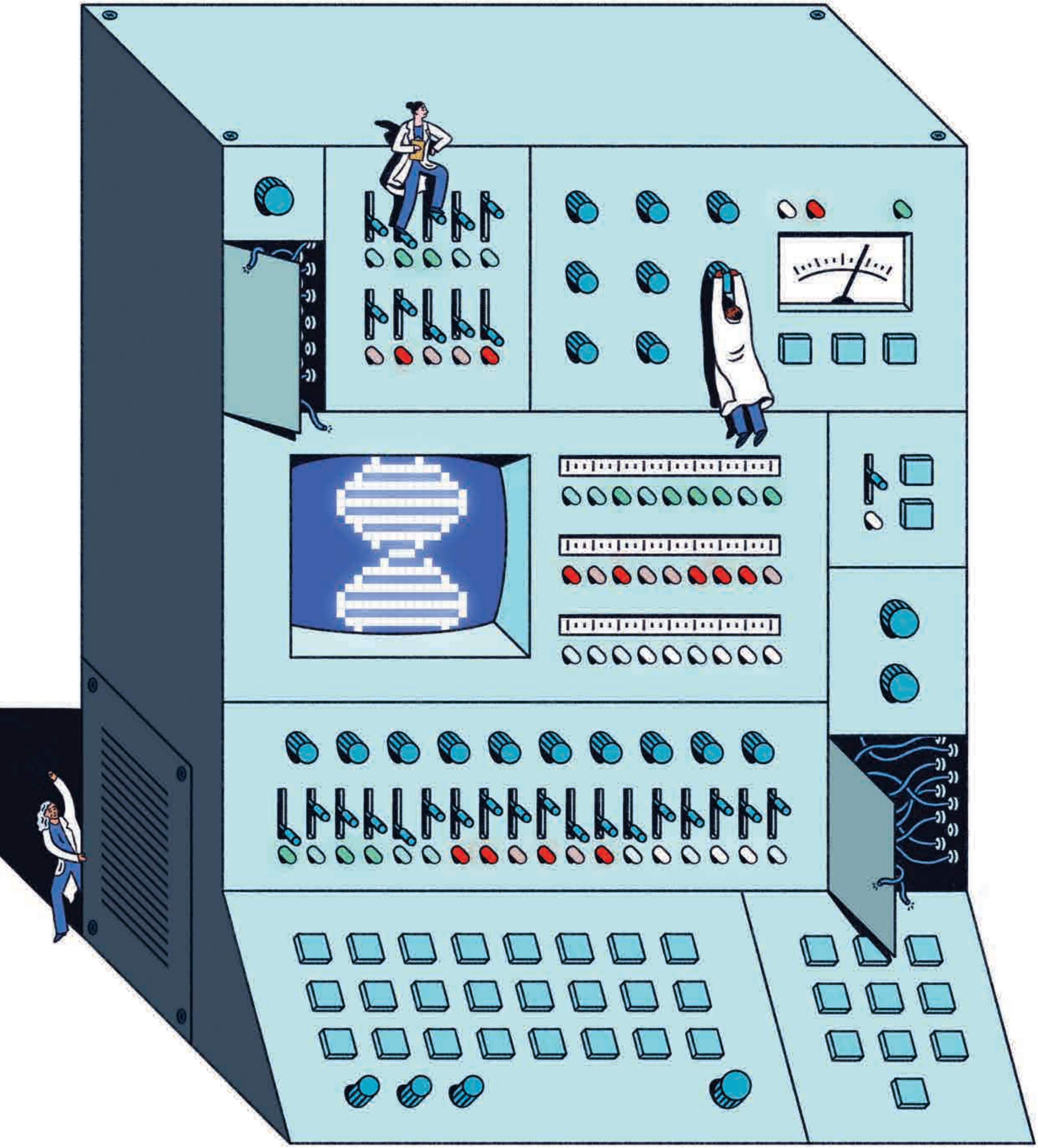


AI COMES OF AGE

Columbia Mailman School researchers are using artificial intelligence (AI) to speed up laborious research steps, probe minute variations in proteins with speed and accuracy, and develop ways to predict disease progression—all while working to ensure that AI tools for public health are both actionable and trustworthy. By Carolyn Wilke/Illustration By Josie Norton

IN AUGUST 2024, EXPERTS FROM ACROSS COLUMBIA UNIVERSITY GATHERED IN COLUMBIA MAILMAN SCHOOL'S HESS COMMONS. AS THEY SAT UNDER WINDOWS OVERLOOKING THE HUDSON RIVER AND THE PALISADES, THE GROUP—COMPUTER SCIENTISTS, BIOSTATISTICIANS, AND ENGINEERS—FOCUSED THEIR ATTENTION ON SOMETHING ELSE: ENVISIONING HOW ARTIFICIAL INTELLIGENCE (AI) COULD USHER IN A NEW ERA FOR PUBLIC HEALTH RESEARCH. Attendees mulled over the limitations of current tools and brainstormed collaborations. The daylong research salon was the first event held by TRAIL4Health, the Translational AI Laboratory for Health, which is tasked with exploring best practices for using AI in public health, building useful tools, and forging partnerships for better research.

TRAIL4Health emerged from the Department of Biostatistics at Columbia Mailman School. "We feel machine learning and AI can be transformative," says Ying Wei, PhD, professor of Biostatistics and the director of TRAIL4Health. Even before ChatGPT's capabilities blew people away, Wei and Kiros Berhane, PhD, the Cynthia and Robert Citrone-Roslyn and Leslie Goldstein Professor and Chair of Biostatistics, were considering how AI might aid research. The potential seemed clear—AI tools are supercharged prediction machines

"THE FUTURE LIES IN EMBRACING AI, BUT IN A WAY THAT DOESN'T DO AWAY WITH THE

that handle huge amounts of data. But, Wei says, those benefits needed to be weighed against the high stakes of public health research. It's one thing for AI to recommend a movie or spit out email drafts. But AI for health needs to be trustworthy, generalizable, and actionable, Wei says.

That's where biostatisticians can contribute their expertise in gauging uncertainty in data, Berhane notes. If researchers can get it right, AI tools may enable researchers to mine unwieldly, underutilized datasets. For instance, wearable devices can rack up 24/7 readings of blood sugar, oxygen levels, and other metrics. Past data science tools used summaries because they couldn't handle the glut of data. Now, "you can deal with the data in its full complexity," he says. That could enable researchers to pursue more complex and important questions.

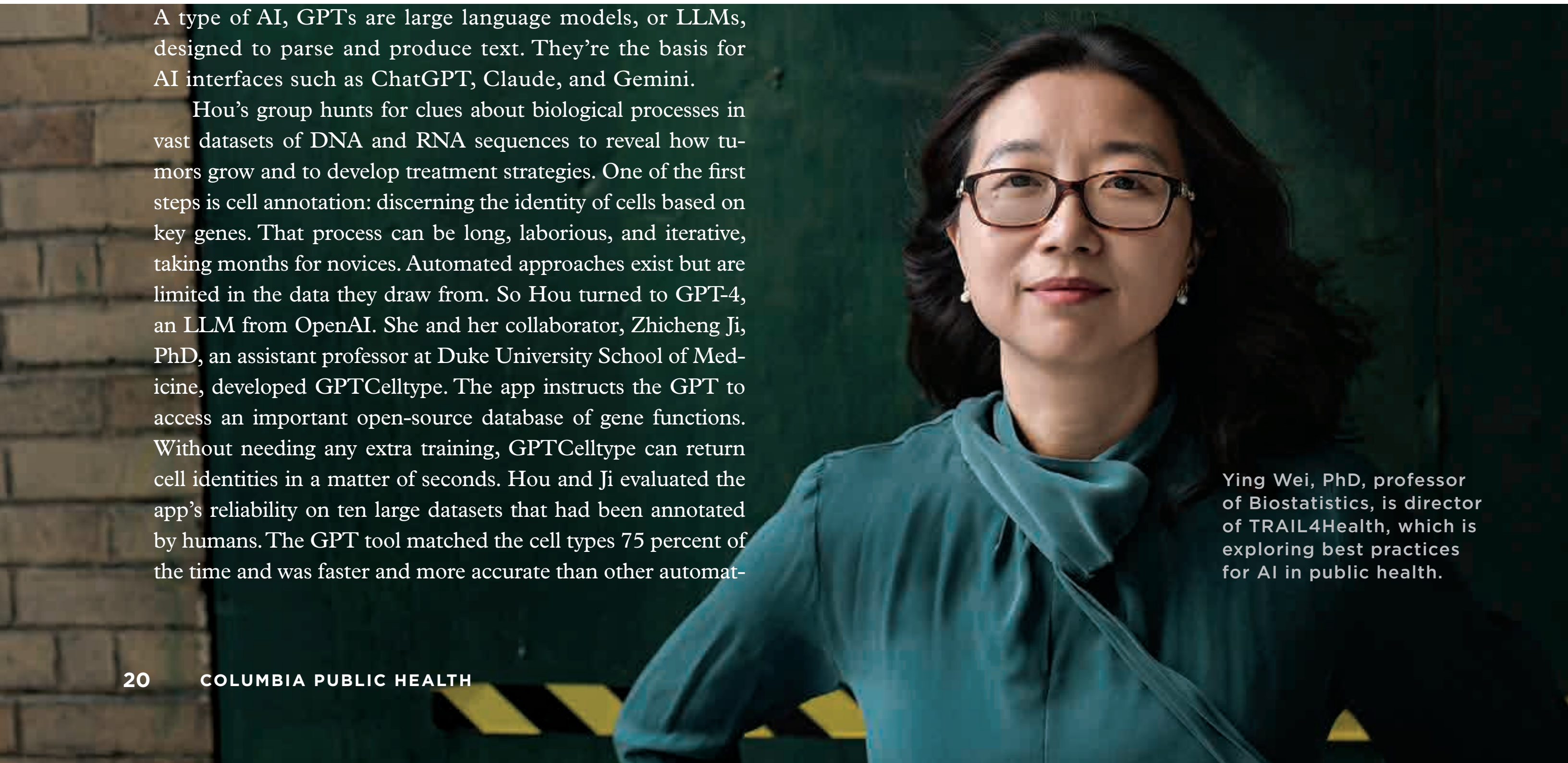
As they continue to interrogate the limitations of AI tools, the School's researchers are using them to make laborious research steps lightning fast and develop ways to predict disease progression. Says Berhane, "To me, the future lies in embracing AI, but in a way that doesn't completely do away with the careful public health research and practice we've been doing so far."

WHEN WENPIN HOU, PhD, ASSISTANT PROFESSOR OF BIOSTATISTICS, STARTED HER LAB IN 2022, SHE SAW STEEP LEARNING CURVES THAT STYMIED SOME TRAINEES. Students and postdocs often struggled with analyzing genomics data, working with biomedical images, or learning programming skills. "GPT models drew my attention quickly when I noticed their great potential to assist these trainees," Hou says. GPT stands for "generative pre-trained transformer," and Hou sensed that the technology might transform learning. But Hou also wondered about the models' reliability. A type of AI, GPTs are large language models, or LLMs, designed to parse and produce text. They're the basis for AI interfaces such as ChatGPT, Claude, and Gemini.

Hou's group hunts for clues about biological processes in vast datasets of DNA and RNA sequences to reveal how tumors grow and to develop treatment strategies. One of the first steps is cell annotation: discerning the identity of cells based on key genes. That process can be long, laborious, and iterative, taking months for novices. Automated approaches exist but are limited in the data they draw from. So Hou turned to GPT-4, an LLM from OpenAI. She and her collaborator, Zhicheng Ji, PhD, an assistant professor at Duke University School of Medicine, developed GPTCelltype. The app instructs the GPT to access an important open-source database of gene functions. Without needing any extra training, GPTCelltype can return cell identities in a matter of seconds. Hou and Ji evaluated the app's reliability on ten large datasets that had been annotated by humans. The GPT tool matched the cell types 75 percent of the time and was faster and more accurate than other automated methods. GPTCelltype quickly began to change how Hou's students work. One master's student flew through the initial step—annotating cells with the tool and verifying results—so he could get right to investigating radiation's effect on human hematopoietic stem and progenitor cells. "Our trainees now can focus on the more intellectual parts of the research," Hou says.

Working with tools that combine text and image processing, Hou has found that GPT models can classify medical images, for instance, categorizing brain MRI scans and lung CT scans as normal or having a tumor. Given one set of labeled training images, the best-performing AI tool identified 89 percent of images correctly and, when prompted, provided the rationale behind its classification. Feeding that rationale back to the AI in a second training boosted image identification accuracy to 99 percent. Language models are an unexplored way to work with images, Hou notes. "This is something really new and inspiring."

Hou's team is also working on an AI-based method to puzzle out epigenetic modifications, chemical changes that alter how cells follow genetic instructions, across different types of cancer and across the genome. The method could reveal hallmarks of specific cancers, making it useful for early detection or identifying new targets for treatments. It's an ambitious project requiring intense computational power, which Hou has access to thanks to financial support from the National Institutes of Health (NIH), IT support from Columbia University, and Google's cloud credits program. The state of New York's Empire AI initiative, which aids public and private research institutions advancing AI research, is helping too. "Access to the high-performance computing resources provided by the Empire AI platform will be critical to the success of this project," Hou says.



Ying Wei, PhD, professor of Biostatistics, is director of TRAIL4Health, which is exploring best practices for AI in public health.

CAREFUL PUBLIC HEALTH RESEARCH AND PRACTICE WE HAVE DONE SO FAR."—BERHANE

AI TOOLS MAY ALSO HELP RESEARCHERS GET A BETTER HANDLE ON ALZHEIMER'S DISEASE. Proteins, produced by cells in response to genetic instructions, have myriad biological effects—from making us itch to allowing us to digest food. And they influence or are influenced by diseases such as Alzheimer's. There are thousands of proteins in blood plasma. Zhonghua Liu, ScD, assistant professor of Biostatistics, reasoned that some of them might foretell the emergence of Alzheimer's well before symptoms set in. His team used a statistical approach to identify seven plasma proteins linked with Alzheimer's disease. They then studied how genetic mutations can change a protein's amino acid sequence, which may alter its three-dimensional structure. "It's like a lock and key," Liu says. Modify the key, and it may lose the ability to open a lock. Changing just one DNA letter in a protein's code can mean the protein loses the ability to participate in a chemical reaction or make a molecular cut.

Liu's team used an AI-based tool called AlphaFold3 to predict folded structures of proteins based on their amino acids. The tool reveals how genetic differences translate to differences in proteins among Alzheimer's patients. Differences in structure can explain why a drug may work for one set of patients but not another. "We need a more personalized treatments," Liu says. Some of the proteins Liu found are already targeted by drugs for other conditions. Clinical trials could reveal whether they hold promise for Alzheimer's. Liu recalls an email from an alumnus sharing that his mother, grandmother, and sister all had Alzheimer's disease. "He hopes someday we can develop a cure for Alzheimer's disease," Liu says. That is what Liu is working on now, and AI is playing a critical role: He is helping to develop a tool to identify moleculer formulas for compounds that will take the right shape to latch onto proteins involved in the disease.

Other researchers, including Wei, are using AI to look for road maps of Alzheimer's disease progression and prevention strategies, with the support of the NIH and the Columbia University Data Science Institute's 2025 Seed Fund. "The question that patients ask the most is, 'What's going to happen next year?'" Wei says. Wei has partnered with colleagues at the Columbia University Alzheimer's Disease Research Center to curate and mine rich patient data—brain scans, blood test results, cognitive assessments and more—with GPT tools. She hopes to uncover which factors predict progression. This could someday help forecast the disease trajectory to help patients prepare, and could also yield insights into preventing or slowing disease progression.

The complexity of the disease and the aging process is a major reason to use AI, Wei says. Simplified Alzheimer's models haven't been adequate for understanding and addressing person-to-person differences. With the amount of data AI tools can integrate, "that could really be transformative." Both Liu's and Wei's approaches could be applied to other diseases, including diabetes, and conditions such as aging. "The modeling capacities could be game-changing," Wei says.

NOT ALL AI APPLICATIONS SUCCEED. Andrew Rundle, MPH '94, DrPH '00, professor of Epidemiology, wondered whether generative AI could boost research efforts in population health. "Somewhat facetiously, the idea was, can we have ChatGPT do the work that we would normally tell doctoral students to do?" he says. In one project, doctoral student Nicole Itzkowitz pulled information about helmet use from anonymized health records of emergency room visits by patients who had bicycle or scooter crashes. These records include doctors' freehand notes. "There's often a ton of information there, but it's a nightmare to use," Rundle says. Working with postdoc Kathryn Burford, PhD, Itzkowitz scoured records for misspellings and abbreviations, eventually creating a text-searching program that categorized reports by helmet use. Could ChatGPT replicate her work? "Yes. On some days," Rundle says.

Fed all the information from the student's painstaking work, the AI tool performed well. But its output wasn't repeatable. Given the prompt and data on successive days, it provided correct information one day and hallucinated the next. On the next day, it repeated the wrong information from the previous hallucination. These experiences have Rundle contemplating how and when AI tools should be used. In some cases, they may shortchange student learning, he says. Meanwhile, Rundle has found his own copyrighted research in databases of articles used for AI training. "Am I going to use this tool that hasn't really produced great results to do my work when it is based upon what I see as the theft of scientists' work?" Rundle says. Given public health's high ethical standards, it's an important question.

How to apply AI tools—with all their capabilities, limitations, and ethical quandaries—is something that Columbia Mailman School faculty, especially those involved in TRAIL-4Health, will continue to consider. Some researchers are developing tools with built-in safeguards to check for consistency and look for errors, Berhane points out. Another issue is AI tools' propensity to propagate biases from training materials or to eliminate or discount data from certain populations. Researchers including Daniel Malinsky, PhD, assistant professor of Biostatistics, are working on measures to counter this algorithmic unfairness. Training—both training AI tools with the best data possible and educating students to interrogate their research results—will be key in using AI well in public health, Berhane notes, but he remains optimistic about the future. "With best practices in place, AI tools will allow us to answer important questions in their full complexity." ●

Carolyn Wilke *covers science for curious readers old and young in* The New York Times, National Geographic, *and more.*

Photograph: John Herr