



Doctors examine images of a patient's lung at a hospital in Xiaogan, Hubei province.

China's data-driven dream to overhaul health care

Collaborations between AI researchers and China's medical workers are helping to combat diseases such as diabetes and COVID-19. **By Sarah O'Meara**

Sitting in her Beijing apartment in early 2020, Guangyu Wang watched as COVID-19 cases increased in China and the rest of the world. The biomedical engineer, who works on artificial intelligence (AI), knew from colleagues working in health care that the country's medical system wasn't coping. The Chinese city of Wuhan had gone into lockdown after recording the first outbreak of COVID-19, and the World Health Organization had declared a global health emergency.

As the number of patients in hospitals grew, so did delays in medical imaging and other

procedures. Chest X-ray scans stacked up, awaiting analysis by radiologists, says Wang, who works at the Beijing University of Posts and Telecommunications. "The health-care system was stretched to the limit," she says.

Wang and others in her team began to reach out to their academic networks – many of whom were keen to use their skills to support the health-care system. Together, they formed a collaboration with biomedical engineers, radiologists, respiratory experts and clinicians from institutions and hospitals across China.

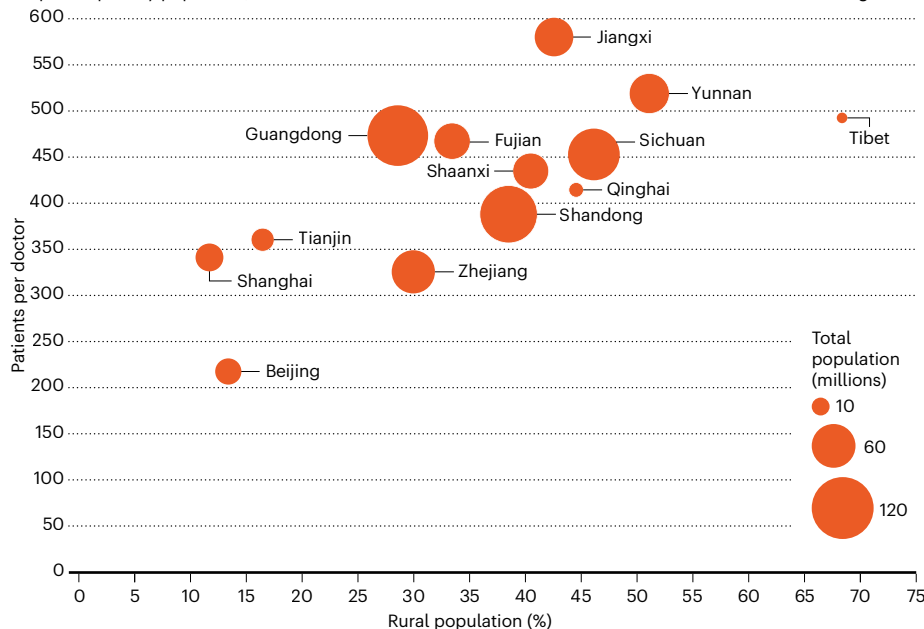
Deciding to focus on medical imaging, the team of specialists worked long days over

two months to develop machine-learning software. By training the software on more than 145,000 chest X-ray images from 7 hospitals, they developed an algorithm that could detect respiratory diseases, including COVID-19, with more than 90% accuracy. The software has since been used to reduce the burden on radiologists working in six hospitals during the pandemic, Wang says. The data and computer code have been deposited with the China National Center for Bioinformation (CNCB), an open-access data platform, to assist the global COVID-19 research effort.

One of Wang's collaborators on the project

MEDICAL ACCESS

People in sparsely populated, rural areas of the Chinese mainland have access to fewer doctors on average.



is respiratory-disease researcher Weimin Li at Sichuan University in Chengdu, who is president of the West China School of Medicine and West China Hospital. He says that, despite coming from a clinical background, he feels comfortable working with AI and informatics researchers. Collaborations like these are increasingly common because, in recent years, government departments such as the National Health Commission, which formulates health-care policies, and the Ministry of Education have actively encouraged clinicians to work with such experts.

A long-term plan

Wang and Li's project is just one result of China's long-term planning and investment in AI technologies, centralized health-care data and a collaborative culture involving researchers and clinicians. Over the past ten years, central government funding and top-down policies have helped to transform Chinese medical research into a data-driven field that uses computer and machine engineering to take pressure off medical workers.

Biomedical engineer Yipeng Hu at University College London says this initiative has led to cooperative, driven Chinese research teams that work at speed. "If something works, it can often quickly be tested and used in hospitals faster than it might be in other countries," he says.

Such enthusiasm can be seen in academia and the private sector, too. "There has been rapid development in AI technologies," says Wang. "The field of intelligent diagnostics is

expanding fast. And China has made a massive investment in the country's health-care research infrastructure."

The numbers bear this out. After China's 12th Five-Year Plan started in 2011, the central government invested more than 10 billion yuan (US\$1.5 billion) in 'health informatization' – the use of information technology to improve health care. (China's five-year plans are published government goals for policy, infrastructure and investment that are revisited every five years.)

"More and more scientists began to believe that AI could be useful in biomedical studies."

There have been further surges in funding since 2010 that have prompted the development of AI research and expertise. The National Natural Science Foundation of China (NSFC), the country's largest funding agency, tripled its budget for research projects from \$4.6 billion in 2006–10 to \$13.7 billion in 2011–15, reaching \$19.6 billion over the period 2016–20.

These funding increases, representing a rise of 17% per year from 2000 to 2017, have closed the gap between China and the United States in research and development (R&D) funding. Preliminary 2019 data from the US National Science Foundation (the most recent data available) suggest that China has surpassed its

technological competitor on R&D spending.

In 2012, for example, China overtook the United States in the number of papers its researchers published in the combined fields of computer science, AI and biomedical engineering. By 2019, that lead had widened, with China publishing around one-third more papers than the United States – although, in the same year, Chinese AI researchers were cited about 20% less than the world average.

Researchers say the investment was sorely needed. Even with increased resources and a rapid publication rate, China has a limited number of highly experienced data engineers and AI researchers who are able to work on projects that combine machine learning and medical expertise. Furthermore, large, high-quality medical data sets can still be hard to come by, says Yu Xue, a computer engineer and cell biologist at Huazhong University of Science and Technology in Wuhan.

Recent improvements to health-data infrastructure have improved matters. A lot of patient data are now stored digitally, and clinical institutions have been instructed by government policies to boost health-care innovation by sharing data more effectively.

However, it remains challenging to translate these new data into clinical applications, says Wang. Any biomedical engineering breakthroughs to come out of such data still require rigorous testing, integration with equipment (which varies between hospitals) and education for clinical staff on how to interpret results.

Hu says that AI researchers worldwide face similar challenges. "Commonly, there is an administrative battle to bring it about. So most research ends up going no further than our research papers, which are based on existing data sets."

Game changer

There has also been a shift in mindset among biomedical researchers towards using AI techniques, says Xue.

As recently as five years ago, many scientists in China were still unconvinced about the power of AI. "There was a deeply held belief that computers could not learn to be intuitive," he says.

Then in 2017, the AI AlphaGo, developed by Google's London-based company DeepMind, defeated the world's number-one Go player Ke Jie – huge news in China, where the ancient board game is wildly popular.

"This tiny storm became a typhoon and blew every scientist's mind," says Xue. "More and more scientists around me began to believe that AI could be useful in biomedical studies."

Xue co-led one project that used deep-learning techniques to develop a software

package that predicts COVID-19 deaths and disease outcomes (W. Ning *et al. Nature Biomed. Eng.* 4, 1197–1207; 2020). He hopes the work can improve the accuracy of disease diagnoses in clinical settings in the future.

Wang has also been involved in developing biomedical AI tools that might aid China's health-care system. She and her colleagues created a smartphone attachment that clinicians and individuals use to take photos of the interior of the eye (K. Zhang *et al. Nature Biomed. Eng.* 5, 533–545; 2021). The images were transmitted to a cloud-computing platform, where neural networks checked them for telltale marks that might indicate a person has chronic kidney disease or type 2 diabetes.

Storage solutions

Xue has advised colleagues involved in setting up two of China's first biological data centres – the CNCB and the National Genomics Data Center (NGDC), both established in 2019. The centres archive, manage and process a wide range of genomics- and other 'omics-related data, serving as China's version of databases such as GenBank in the United States and the European Nucleotide Archive, which were set up in the early 1980s.

"Lack of money prevented China from establishing biological data centres," says Xue. "Now money is no longer a problem, China is looking to support worldwide initiatives to collect, share and distribute data, which is a costly and labour-intensive process."

For example, in January, the NGDC and CNCB launched a database containing genome, nucleotide and protein sequences for the coronavirus SARS-CoV-2. It also includes sequences of other coronaviruses, epidemic data, genome variations and lineage data, as well as clinical records (<https://ngdc.cncb.ac.cn/ncov>). "It's been widely used across the world, and frequently and extensively updated," Xue says.

Data-driven diagnostics

China's government, researchers and clinicians have a strong motive beyond the pandemic to translate computer diagnostics into clinical settings. There are too few qualified doctors in remote areas (see 'Medical access'), forcing people to travel lengthy distances to cities and wait in long queues, putting extra strain on hospital staff and clinical resources.

Li thinks that diagnostic technologies driven by deep learning might offer much-needed support. If people can be diagnosed without the direct involvement of a time-poor doctor, there might be no need to visit a hospital. Basic treatments such as antibiotics or painkillers could also be provided locally, he says.



Artificial intelligence developed by Hefei-based technology firm iFLYTEK and the Chinese Academy of Sciences can analyse patients' imaging scans.

Such advances are already being implemented. One of China's largest AI and speech-technology companies, iFLYTEK in Hefei, has developed an AI Medical Assistant that is in use nationwide. Health-care staff enter a patient's symptoms into the software platform, and it provides advice on a possible diagnosis. According to the company, the platform gives around 400,000 of these auxiliary opinions each day in as many as 30,000 medical institutions.

Many of China's technology giants, including Alibaba, Tencent and Baidu, have announced various AI-driven health-care diagnostic tools in the past few years. Baidu, for instance, used an AI-powered mapping platform to track the movement of people leaving Wuhan as the first cases of COVID-19 were identified, which it says was used to "accelerate local preparedness and response efforts".

There is also a funding incentive for researchers. In November 2020, the NSFC established a department to support collaborative research across four key areas: broadly, the physical sciences; information technology and AI; biomedicine; and the humanities. It will be allocated a proportion of the total budget spend from the NSFC, which in 2020 was 28.5 billion yuan. "I would estimate it might get 5–10% of all funding," Xue says – or up to 2.8 billion yuan.

Tricky translation

However, translating research into clinical tools is still far from straightforward and depends on the availability of huge, well organized and annotated data sets. Not every medical institute has the capacity to provide these, says Xue.

There are also issues around data protection.

In China, a legal framework came into effect on 1 September that aims to further strengthen legal protections for Chinese citizens' personal information.

Overall, the guiding objective of China's system is similar to that of Europe's General Data Protection Regulation laws, says Rebecca Arcesati, a technology and digital-policy analyst at the Mercator Institute for China Studies in Berlin. Regulators have tried to balance personal privacy, economic development and national-security considerations, she says.

"The law is the Chinese government's response to long-standing public concerns over privacy violations by powerful tech giants, who so far have operated and grown in a de facto Wild West," she says.

However, the framework differs from European legislation in how it treats personal information: considering it as a strategic resource that belongs to the Chinese state before its citizens, she says. "What is important to stress, in my view, is that the law is unlikely to limit the widespread forced collection of biometric data and genetic information from the population, especially from vulnerable minorities."

Given China's growing economy and its ability to pay competitive wages to talented early-career AI researchers, Hu thinks that it won't be long before the country will match the innovation levels of its rivals in Europe and the United States, in terms of citation rates and breakthroughs such as those achieved by AlphaGo. Young Chinese researchers learn quickly, says Hu. "The gap in their experience, which limits their ability to be truly innovative, could be reduced very quickly in the future."

Sarah O'Meara is a freelance journalist based in London.