

Issue Brief

China and Medical AI

Implications of Big Biodata
for the Bioeconomy

Authors

Caroline Schuerger

Vikram Venkatram

Katherine Quinn

Executive Summary

Medical artificial intelligence is an area of emerging technology that leverages biomedical advances and AI to power new discoveries for human health. Medical AI can help researchers discover new medicines, understand complex causes of disease, and predict how patients will respond to therapies. As these applications improve public health outcomes, they also contribute to the growing global bioeconomy. Countries that strategically prioritize medical AI could enjoy a competitive economic advantage and set global standards and norms for future developments.

Medical AI leadership depends on access to large repositories of biological data, or biodata. Biodata can include information collected from individuals, such as their unique genetic sequences or medical records. Researchers and developers who have access to large amounts of high-quality biodata are at an advantage, as this resource is a limiting factor for medical AI advancement.

China has created a comprehensive national strategy to support medical AI development and advance its goals for bioeconomy leadership. This report examines China's stated goals for medical AI, which range from the collection and protection of vast amounts of biodata, to facilitating research and development, to supporting medical AI commercialization. We find that:

- **China has access to publicly-available biodata from around the world, while its domestic datasets are closed off to other countries.** Access to more biodata facilitates more powerful medical AI applications.
- **Medical AI research publications from both China and the United States are on the rise.** The United States only narrowly led China in the production of research publications related to medical AI in 2021, and China's research output is likely to continue to grow due to the many resources the Chinese government is pouring into this area.
- **Beijing is promoting medical AI commercialization, and top Chinese technology companies that have not previously been involved in biotechnology are moving into the medical AI space.** Policies to support companies and facilitate regulatory approval can accelerate medical device development.

In sum, China's strategy for biodata collection and medical AI development could position it to be an economic and technological leader in this sector. Policymakers should consider the implications of this leadership when considering policies to boost U.S. competitiveness and biodata infrastructure.

Table of Contents

| | |
|---|----|
| Executive Summary | 1 |
| Introduction | 3 |
| Biological Data: A Societal Resource Fueling AI Innovation | 4 |
| Policies to Make China a Global Biodata Leader..... | 9 |
| Policies Facilitating Biodata Accumulation..... | 9 |
| Policies Controlling Access to Data | 11 |
| Policies to Support Research, Development, and Technology Infrastructure | 13 |
| China’s Medical AI Research Output | 13 |
| Chinese Companies Developing Biodata-Enabled Medical AI..... | 16 |
| Conclusions and Lasting Implications of China’s Biodata Leadership..... | 19 |
| Appendix A: Select List of China’s Notable Governmental Policies on Medical AI..... | 20 |
| Authors..... | 24 |
| Acknowledgments..... | 24 |
| Endnotes | 25 |

Introduction

As science and technology advance, cutting-edge developments in artificial intelligence are accelerating biological innovation. AI allows researchers to process enormous amounts of complex data more quickly and efficiently, and can advance health and medical applications by identifying patterns in large repositories of biological data (hereinafter referred to as biodata). For instance, during the COVID-19 pandemic, countries around the world used AI tools to diagnose disease, predict therapeutic responses, and calculate infection probabilities.¹

Looking ahead, there will be huge economic and societal rewards for countries that lead the development and deployment of AI technologies for human health purposes (hereinafter referred to as medical AI). The global bioeconomy is rapidly growing; some experts estimate that bio-produced products could contribute up to \$4 trillion per year to the global economy by the end of this decade, and that medical AI alone will be worth over \$12 billion globally by 2030.² In addition to economic power, the leader in medical AI can set global standards for research, technology deployment, and data usage. Given these potential advantages, countries have an incentive to invest resources and create dedicated policies to advance the technological development of medical AI, including by making high-quality biodata repositories more accessible.³

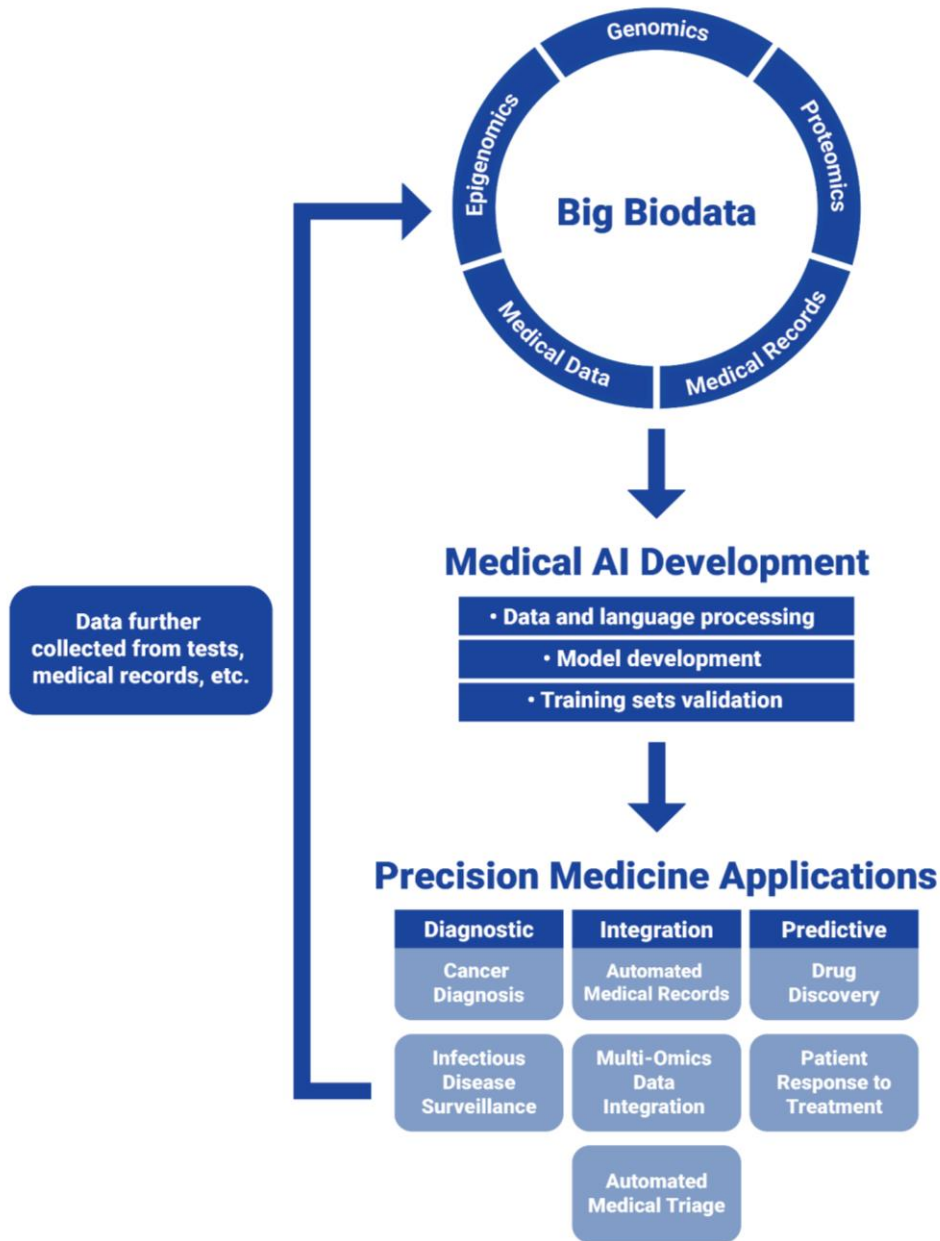
China is one of the best examples of a country that has created a strategy to integrate biotechnology and AI. China can access publicly-available data from around the world, while maintaining domestic datasets that are closed off to other countries.⁴ Many Chinese government policies explicitly state their intent to prioritize biodata collection, application, and medical AI deployment, as do statements from China's medical AI industry.⁵ While we analyze those published declarations and strategies, their implementation and eventual success remain to be seen.

This brief examines China's stated and understood goals for biodata and medical AI in a global context, focusing on the policies and programs the Chinese government has put in place to collect and analyze vast amounts of biodata including genetic, proteomic, and clinical data. It also examines different aspects of the ecosystem that China is developing for cutting-edge technologies, including medical AI, in order to compete and dominate this growing aspect of the bioeconomy. This brief concludes by assessing the implications of these developments for U.S. competitiveness in this sector. It is important to note that, while this report acknowledges the inherent national security and ethical implications associated with China's efforts surrounding biodata, it focuses primarily on the economic implications of these efforts.

Biological Data: A Societal Resource Fueling AI Innovation

Biological data, or biodata, is a broad category that encompasses large datasets of biological, life sciences, or biomedical information. Among other examples, biodata can include information collected from individuals, such as their unique genetic sequences or medical records (types of biodata summarized in Table 1). AI technologies can make use of this huge amount of data to identify links between genes, diseases, and treatments that would otherwise be difficult and time-consuming to recognize. Cutting-edge medical AI applications are increasingly being adopted in fields like personalized medicine, drug development, and disease surveillance. As these tools are implemented, their use generates and collects even more biodata that can feed into the next cycle of development (Figure 1).

Figure 1. Cycle of AI-Enabled Medical Technology Development



Source: CSET analysis.

Increased access to biodata can support and accelerate basic research, and help scientists and doctors analyze demographic health trends or such trends across the general population.⁶ That said, making efficient use of biodata and particularly DNA data resources can be challenging and time consuming. This is in part because most biodata datasets are large; for example, there are about three billion nucleotide base pairs that make up about 20,000 genes in a single human's genome.⁷ New AI tools, though, can significantly shorten the time it takes to process, investigate, and provide insights into these large datasets.

A particularly promising area where AI is already being applied to analyze biodata is precision medicine. A patient can respond very differently to various treatments based on their specific medical history, physical characteristics, and genetic makeup. Precision medicine takes these differences into account, using a patient's unique biodata to help decide which treatments and therapies are likely to work best for that individual.⁸ For example, knowing the DNA sequence of a patient's cancer cells can indicate which genetic mutations led to the disease, and help doctors choose between treatment options that are personalized to the patient's specific diagnosis.

Overall, medical AI is of great value to precision medicine and public health across a broad range of applications, including:

- Algorithms to predict how a patient will respond to a specific treatment.
- AI-guided drug discovery.
- Integrating electronic health records into genetic databases to connect specific DNA mutations to disease outcomes.
- Using algorithms trained on medical images, such as MRIs or X-rays, to aid doctors in diagnosing patients.
- Using data from medical records to track the spread of infectious diseases.

Table 1: Types of Biodata Used for Developing Next-Generation Technology Applications in the Medical Field

| Data | Example(s) | Definition and Function |
|----------------------------|---|---|
| Genomics | DNA sequences | The human genome is made up of 3 billion base pairs of DNA, some of which encode genes, the instructions for making proteins with specific functions. Mutations, or changes to a gene, can cause a disease or affect how the body works. ⁹ An individual's genetic makeup can impact their response to certain medications, or their likelihood of developing a certain disease. |
| Epigenomics | Gene expression, DNA or histone modifications | Epigenetics refers to reversible alterations that change how and when certain genes are expressed, or converted to proteins. Typically, epigenetic changes refer to chemically altering DNA or how it is packaged, but do not change a gene's DNA sequence. ¹⁰ |
| Proteomics | Protein sequences, functions, and 3D structures | Proteins are functional components that perform jobs within cells. Proteins are encoded by genes, which must be built, or expressed. Proteomics studies protein characteristics, presence, location, interactions with other proteins or cellular components, etc. ¹¹ |
| Medical Images | X-Rays, CT scans, MRIs, ultrasounds | Medical images capture visual and electromagnetic data from patients. ¹² This information can help to diagnose diseases or medical conditions, or to track how a disease progresses over time. |
| Electronic Medical Records | Patient medical history | Electronic medical records, or electronic health records, include a patient's specific health information like their biometric or demographic information, disease status, and medication history. This data gives insight into the various, complex factors that make up a patient's health status. ¹³ |

Source: CSET analysis.

Medical AI's potential to meet public health needs and overcome cost, labor, and access challenges has many countries striving to harness AI's capabilities in the medical field.¹⁴ Many technologically advanced countries around the world, including the United States, the United Kingdom, and European Union, are prioritizing medical AI development through mechanisms including research funding, database maintenance and curation, and collaboration with the private sector.¹⁵ For example, U.S. Executive Order 14081 on Advancing Biotechnology and Biomanufacturing Innovation cites the need to “foster a biological data ecosystem that advances biotechnology and biomanufacturing innovation” to “drive breakthroughs for the United States bioeconomy.”¹⁶ While there is a worldwide effort to advance medical AI, China's efforts are particularly notable seeing how its government has adopted wide-ranging policies for collection and protection of biodata that could give it an advantage in furthering its broader bioeconomy and AI goals.

Policies to Make China a Global Biodata Leader

The Chinese government is prioritizing the collection, protection, and promotion of biodata to advance its goals for bioeconomy leadership. While many countries have created strategies to promote AI for public health and precision medicine, medical AI could be especially beneficial for China by solving both urgent and growing public health needs brought on by the country's large and aging population, as well as its disparate rural healthcare system.¹⁷ China has turned to biodata to develop medical AI for major public health applications like disease surveillance, vaccine development, and food security issues.¹⁸ For example, AI tools can reduce the number of physicians needed by helping doctors to diagnose, triage, and predict the best treatment outcomes for patients.¹⁹ Tools like digital health services, which integrate medical and clinical information on cloud servers, can help patients in rural areas that might not have access to big medical facilities by reducing in-person hospital visits and streamlining data sharing for physicians to make clinical decisions.²⁰

The Chinese government's central policy planning documents and regulations have consistently prioritized biomedical sciences and AI innovation to pave the way for medical AI. In particular, the Chinese government's commitment to leverage biodata to advance precision medicine, health, and AI is highlighted in the Five-Year Plans and the National Medium- and Long-Term Plan for Science and Technology Development (国家中长期科学和技术发展规划) as well as in major health-related initiatives like Beijing's Healthy China 2030 Policy (“健康中国2030”规划) (see Appendix A for a table of related Chinese policies, regulations, and guidelines).²¹ The Chinese government's inclusion of medical AI development in general country-wide policy documents indicates that precision medicine and emerging medical technologies are a strategic priority.²² That said, there are U.S. initiatives, like the Precision Medicine Initiative (now the National Institutes of Health's “All of Us” program), that similarly aim to support medical AI development.²³

The following section examines additional top-down, centralized policies and infrastructure Beijing is putting in place to realize these objectives. In addition to the broad priority-setting policies mentioned above, China's efforts break down into two categories: 1) laws and activities facilitating the collection of data, and 2) laws and activities controlling the access to that data.

Policies Facilitating Biodata Accumulation

Over the past decade, China has accumulated massive amounts of biodata through a variety of efforts, including instituting requirements to collect and maintain people's

biodata in big national data centers, digitizing health records, and launching programs to collect biodata from citizens. See Appendix A for a detailed table of many of these policies.

Chinese policies that standardize biodata digitization from physical files have facilitated the creation of large datasets that can then be used for research and development. China's 14th Five-Year Plan for National Informatization (“十四五”国家信息化规划) is just one of several policies that require hospitals to digitize health records and to provide universal digital healthcare, driving further collection of biodata (see Appendix A).²⁴ Additionally, several policies, like Healthy China 2030 (“健康中国2030”规划), instruct standardization of data and connection between data centers, making it easier to use across different research applications and specifically directing data collection to promote big data-informed health applications.²⁵ China's Biosecurity Law (生物安全法) emphasizes the government's sovereignty over biodata, and the country's Measures for the Management of Scientific Data (科学数据管理办法) policy requires that scientific data generated in China considered relevant to national security issues be consolidated in newly developed scientific data centers.²⁶

While some of these measures are not dissimilar from the initiatives countries like the United States and United Kingdom have adopted to promote biodata collection and consolidation, the Chinese government additionally accumulates biodata through initiatives that may not be considered ethical in other countries and are in fact not standard state uses of biodata, especially in the United States, due to both ethical concerns and stringent privacy laws.²⁷ According to Human Rights Watch and the Australian Strategic Policy Institute, Chinese government programs have systematically collected millions of people's biological data, including DNA and blood types, fingerprints, and facial and iris scans, without informed consent or evidence of criminal activity justifying this collection.²⁸ Programs in Xinjiang and Tibet, for instance, have used alleged anti-terrorism campaigns and medical programs as pretext for collecting biometric data.²⁹ Data from these collection methodologies have been traced to research studies: In 2021, the scientific journals *Human Genetics* and *the International Journal of Legal Medicine* retracted two scientific publications from Chinese government laboratories which used DNA collected from Uyghur people in Xinjiang and connected genetics to physical appearances, citing “ethical concerns on collection of data.”³⁰

Beijing is focused not only on collecting domestic biodata, but also supports efforts to gain access to biodata from other countries. One example that illustrates China's central policy initiatives to accumulate biodata from outside of China is BGI (华大基因), formally the Beijing Genomics Institute, a major genetic sequencing company based in

Shenzhen. In China, BGI runs the government-owned China National GeneBank (CNGB; 国家基因库).³¹ But subsidiaries of BGI, such as MGI (华大智造), have popped up around the globe.³² In 2012, MGI acquired the California-based company Complete Genomics, which operates DNA sequencing services in direct competition with the U.S. company Illumina.³³ Another BGI subsidiary, TECH SOLUTIONS (HONGKONG) CO., LIMITED, is certified by the U.S. government to provide laboratory testing on patient samples from U.S. hospitals and medical facilities.³⁴

The fact that BGI has so many avenues of genetic data collection, across China as well as from other countries through the aforementioned subsidiaries, is concerning for a number of reasons. For one, there are reports that the company holds onto data collected from their services and, on certain occasions, has also shared biodata collected from non-invasive prenatal tests conducted inside and outside China with China's government-funded National GeneBank.³⁵ Moreover, according to media reports, BGI has collaborated frequently with the People's Liberation Army, including to develop those prenatal tests, and the PLA has used at least BGI's domestic Chinese genetic data for military research.³⁶

Policies Controlling Access to Data

In tandem with policies designed to promote the large-scale collection of biodata, China is also consolidating its biodata resources by implementing policies and regulations that enhance domestic access and prohibit data from leaving China.

Chinese policies requiring the digitization of health records, like the 14th Five-Year Plan for National Informatization (“十四五”国家信息化规划) mentioned above, aim to make biodata available for researchers in addition to driving more data collection.³⁷ The CNGB, which contains over 10 million genetic data samples, and the National Genomics Data Center (NGDC 国家基因组科学数据中心) allow researchers in China to access data that can be used to develop medical AI tools.³⁸ This practice is not uncommon and has been implemented in other countries; for example, the United States' National Institutes of Health (NIH) National Center for Biotechnology Information (NCBI) and the UK Biobank are sources of government-funded and maintained biodata databases for researchers to use.³⁹ However, while researchers in China can access the world's data through these public databases, the world cannot access China's data due to Chinese data privacy laws that prevent foreign researchers from accessing Chinese databases.⁴⁰ In contrast, U.S. databases like NCBI's GenBank and the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) do not restrict the access, use, or distribution of data.⁴¹

China's companies also benefit from the country's biodata holdings and data standardization policies that make it easier for them to access and use that biodata. In contrast, companies in other countries do not have access to biodata to develop technologies in the same manner or at the same scale. In the United States, for example, companies must partner with different hospitals or research institutions to gain access to certain types of biodata for developing medical AI tools, and cannot access this data on their own.⁴² Moreover, U.S. companies face additional challenges because available biodata is often not standardized across different institutions and public health systems, which delays the development of medical AI tools.⁴³

China's privacy and security laws directly state that foreign entities cannot use Chinese genetic data without approval, or export data outside of China, specifically prohibiting institutions and hospitals with any foreign owners or investors from collecting or preserving any human genetic resource from people within China, including data involved in medical device approval processes (see Appendix A). China's Outbound Data Transfer Security Assessment Measures (数据出境安全评估办法) require that researchers' data be assessed by the provincial-level and national cybersecurity and informatization departments before it is published in an academic research paper in a foreign journal.⁴⁴ Already, these restrictions have forced major scientific publications to grapple with data availability and its relation to transparency and reproducibility. In 2020, for instance, the scientific journal *Cell* published a study even though the authors, restricted by regulations from the Human Genetic Resource Administration of China, refused to deposit the raw genetic data they used for the research.⁴⁵ By contrast, outside of patient privacy laws like HIPAA, there are no similar restrictions preventing data collected in the United States from being published in foreign journals.⁴⁶

The above discussion illustrates that the Chinese government is engaged in a range of efforts to collect massive amounts of biodata, including through unethical means that violate patient privacy as well as by supporting companies that acquire international biodata, as seen with BGI. The Chinese government is also implementing laws that limit access to Chinese data and prohibit data from leaving the country's borders. This means that some data collected in China cannot be used to improve technology in other countries or be shared with the international scientific community.

The following section elaborates on how China is building an ecosystem to facilitate the use of this data, including for medical AI applications. Taken together, these measures could give China a competitive advantage in developing new medical AI technologies, which could boost its bioeconomy in comparison to the United States.

Policies to Support Research, Development, and Technology Infrastructure

To reach its ambitions in biotechnology and AI, Beijing has drafted specific AI-development plans that cross technical areas. These plans urge universities to incorporate AI into biological research and medical healthcare to further basic research in these areas. The 13th Five-Year Plan for the Development of Strategic Emerging Industries (“十三五”国家战略性新兴产业发展规划) specifically called for co-locating universities and institutes in industrial clusters to “accelerate the pace of innovation and development in the biotech industry and foster new biotech economic drivers.”⁴⁷ The AI Innovation Action Plan for Institutions of Higher Education (高等学校人工智能创新行动计划) states that universities should integrate AI into intelligent medical care and promote biotechnology integration with information technology.⁴⁸ While these and other government directives are important to note and track, we must also acknowledge that it is difficult to ascertain their levels of success at this point.⁴⁹

In addition to supporting academic and government researchers, Beijing also provides financial and other support to private companies as part of its broader efforts to accelerate medical AI development. For example, in 2010, China Development Bank (国家开发银行), a state-funded and state-owned bank under the jurisdiction of China’s State Council, provided \$1.5 billion to BGI to extend its research and development platforms.⁵⁰ More recently, China’s biomedical industrial clusters, mentioned above, also co-locate private companies alongside the university programs and government research facilities, to support their research.⁵¹

Beyond support for research and development, China is also amending some of its medical device policies to accelerate the commercialization of medical AI technologies. For example, the National Medical Products Administration (NMPA; 国家药品监督管理局) recently reduced barriers to using biodata to develop medical devices by expanding the types of data that can be used to develop technologies.⁵² This new guidance also reduced certification processing time and allows devices to be certified for longer.⁵³ The NMPA regulation opens the door for companies to use a wider variety of health data in their AI technologies, and reduces bureaucratic barriers to apply for approval certification.

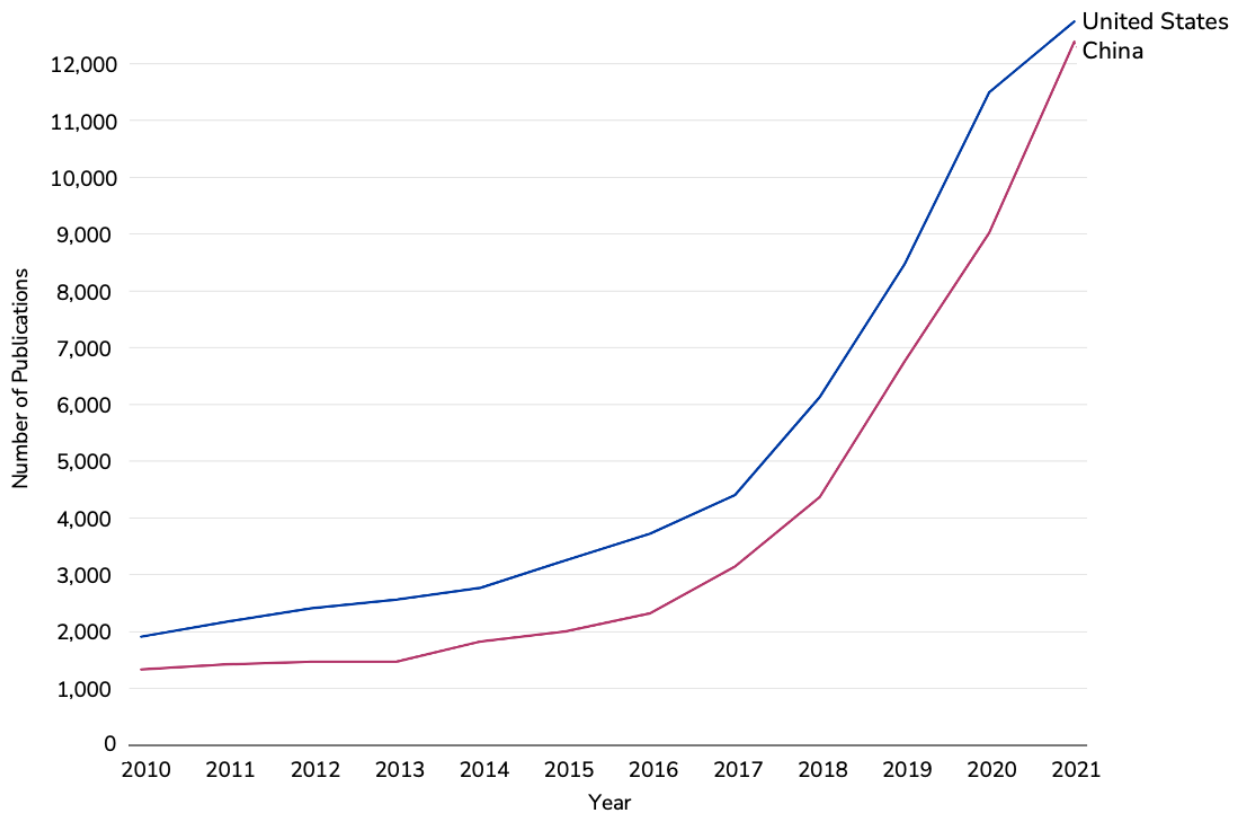
China’s Medical AI Research Output

Over the past decade, China has nearly matched the United States in the number of medical AI-related English-language research publications per year (Figure 2). In this section we take a closer look at China’s medical AI research publications over the past decade, compared to the research output from scientists in American institutions.

Notably, examining publication counts alone does not fully account for a country's innovative capacity and technological development. The sheer number of papers does not account for the quality or impact of those papers, and so the country with the most publications may not be the country making the most influential discoveries. Nevertheless, an assessment of research publication outputs does provide insight into the state of China's medical AI research ecosystem and how it has changed and grown over time, partly as a result of the aforementioned biodata policies the Chinese government is promoting as part of its broader strategy to strengthen the country's bioeconomy.

We drew on CSET's merged corpus of scholarly literature to measure the United States' and China's predominantly English-language research output in medical AI fields by querying keywords that cover both medical and AI topics between 2010 and 2021.⁵⁴ We found that medical AI research publications from both countries grew steadily over this period, with the United States' just barely still leading China by 2021 (Figure 2).⁵⁵ Given the resources the Chinese government is pouring into this area, China's medical AI research output is likely to continue growing.

Figure 2. Trends in Publication Output Containing Medical and AI Keywords for China and the United States, 2010-2021



Source: CSET's merged corpus.*

* For more information on how we generated our merged corpus of scholarly literature, see <https://eto.tech/dataset-docs/mac/>.

Chinese Companies Developing Biodata-Enabled Medical AI

China has been an early adopter of medical AI technology, bolstered by its extensive biodata collection and aggregation efforts.⁵⁶ The NMPA has granted several Chinese companies, ranging from medical to technology companies, approvals to use biodata from many sources to develop and deploy medical AI technologies.

Below are some examples of Chinese companies that received the NMPA's approval to use patient biodata to develop medical AI technologies. These companies report a range of successful use cases where they've deployed medical AI. These claims cannot be taken at face value, since companies have a vested interest in projecting success while not advertising their proprietary capabilities, while China's government also benefits from the positive advertising showcasing the country's biomedical industry. The statements of these companies alone do not prove the success of China's medical AI industry. However, the claims mentioned below do provide evidence that companies supported by and working with the Chinese government are using medical AI systems beyond a research setting, which has implications for the bioeconomy worthy of U.S. policymakers' attention.

- Yidu Tech (医渡科技): According to the company's website, Yidu's AI platforms aim to help China's hospitals operate more efficiently in medical research, medical management, drug research, and clinical decision support. As of 2021, Yidu claims to have analyzed 1.3 billion medical records across 500 hospitals in China, used AI to help diagnose blood diseases, published over 100 papers using data from their platforms' use in hospitals, and provided technology to the Chinese Center for Disease Control and Prevention to assist in pandemic control, assessing transmission risk, and strategizing response.⁵⁷
- Airdoc (鹰瞳科技): The company Airdoc claims to use images of retinas to diagnose various eye diseases, and its AI-based software to diagnose diabetic retinopathy is NMPA-approved. The Airdoc algorithms were trained and developed using over 200,000 retinal photographs from 16 clinical settings in China. The company's founder claims this technology has a higher accuracy rate and allows for quicker diagnoses than doctors who rely on alternative tools can provide, and that it is helping overcome eye doctor shortages in China.⁵⁸
- Infervision (推想科技): Infervision claims to use AI-enabled radiology technology to diagnose cancer through CT scans of lungs, and its InferRead Lung CT.AI technology has received approval for clinical use in the United States and Europe in addition to China.⁵⁹ The company's founder has stated that this

technology will address a major gap in China, as “In China there are just 80,000 radiologists who have to work through 1.4 billion radiology scans every year.”⁶⁰ According to company statements, Infervision developed its technology using stored electronic health data to train their algorithms, and incorporated data from the 20 Chinese hospitals to which it was initially deployed. During the COVID-19 pandemic, Infervision translated their lung scan technologies to identify clinical features of COVID-19 and aid in diagnosing patients.⁶¹

Top Chinese technology companies, including those that have not previously been involved in the medical field, are also moving into the medical device space, which further highlights the market potential of successfully translating biodata into cutting-edge medical technologies. Some examples of this emerging trend include:

- Baidu: According to company reporting, Baidu’s PaddlePaddle (飞桨) AI platform is used in many medical applications from diagnostics to predictions. Doctors use the PaddlePaddle platform to analyze CT scans and detect image patterns to aid in diagnosing pneumonia patients, and most recently, scan the lungs of coronavirus patients for disease. Researchers, including U.S. researchers, used Baidu’s platform to develop an algorithm that predicts RNA folding structures.⁶² Baidu also states that their platform was used during the COVID-19 pandemic to power predictions of coronavirus RNA structure for future pharmaceutical applications.⁶³
- Alibaba: In addition to working with BGI to provide computing power and storage for massive amounts of genetic data, Alibaba also developed their ET Medical Brain (ET医疗大脑) AI platform. The company claims the platform can be used to integrate hospital logistics and diagnostic assistance for better patient outcomes by incorporating data from patient electronic health records, and patient images.⁶⁴
- Tencent: After being tapped by the Ministry of Science and Technology (MOST) in 2017 as a “National AI Team” member, earning government support and access to regional projects and data resources to construct AI platforms for medical imaging, Tencent announced that they launched the Jarvis Lab (腾讯天衍研究中心) in 2018 to focus on developing Medical Artificial Intelligence. According to Tencent, the AI products from this laboratory are used to integrate electronic medical records and disease risk prediction, and automatically analyze medical images such as CT and PET scans to aid doctors in diagnostics. Tencent has also reported a partnership with pharmaceutical giant Novartis on an AI

Nurse platform for digital health access in China, and claims to be working to make their AI-powered digital health service WeDoctor (微医) public.⁶⁵

The companies involved in medical AI development and deployment have also faced some challenges. For instance, Baidu, Alibaba, and Tencent (BAT) have reportedly had trouble gaining regulatory approval from the NMPA, unlike many smaller players in the medical AI space who have had more success due to their increased strategic focus on medical AI compared to BAT's more disparate priorities.⁶⁶ However, the efforts of a variety of Chinese companies to enter this industry reflect China's push towards medical AI.

Conclusions and Implications of China's Biodata Leadership

China is making rapid progress in developing and implementing medical AI technologies, in part because of national strategies that facilitate the collection and integration of biodata. Beijing's dozens of dedicated policies to collect and use health data demonstrate its prioritization of this issue, and are likely to lead to continued growth in medical AI research. As China develops these tools and brings them to international markets, U.S. policymakers will have to contend with various challenges.

In particular, it is important to consider how China's accumulation and use of biodata impacts the United States' competitiveness in the global bioeconomy. Chinese researchers have access to far more biodata than U.S. researchers do, including large public datasets from the United States, Europe, and others, plus all of the genetic data accumulated within China that U.S. researchers do not have access to. While nothing is for certain, this comparative advantage could help China become a global powerhouse in medical AI and capture the economic value that comes with it.

Furthermore, whoever dominates the space of next-generation medical technologies will have an outsized impact on research standards and norms, and on how future developments and applications are prioritized. If China becomes the world leader in this arena, future developments may not align with U.S. priorities due to differences in research transparency, privacy, and medical device approval mechanisms. Major Chinese technology champions are already permeating the global health space; for example, Infervision's "InferRead Lung CT.AI" tool is approved as a medical device in the United States by the U.S. Food and Drug Administration.⁶⁷ When medical AI technologies from Chinese companies are approved and adopted in the United States, health data from U.S. patients could be collected and fed back into China's databases. While this approach to biodata collection and accumulation could lead to more efficient and beneficial medical AI, it would also widen the data resource gap between China and the rest of the world.

In summary, the Chinese government has created a unique ecosystem that treats biodata as a resource, and accumulates and uses it to advance medical AI development. This centralized strategy may give China an edge against the United States, and its significant economic, societal, and technological implications should be considered in future policies that aim to bolster the U.S. bioeconomy.

Appendix A: Select List of China's Notable Governmental Policies on Medical AI

| Policy | Year | Implications |
|--|------|---|
| Guidelines for the Registration and Examination of AI Medical Devices (人工智能医疗器械注册审查指导原则) ⁶⁸ | 2022 | <ul style="list-style-type: none"> Introduces a more detailed definition of allowed data sources for AI devices, expanding the scope from “data from medical devices” to encompass medical and non-medical information, including medical records, exam results, and patient complaints, etc. Expedites process and certification length. |
| Outbound Data Transfer Security Assessment Measures (数据出境安全评估办法) ⁶⁹ | 2022 | <ul style="list-style-type: none"> Requires oversight and restricts data of any kind leaving China by conducting security assessments. Requires oversight of any researcher's data before using it in an academic research paper in a foreign journal. |
| 14th Five-Year Plan for National Economic and Social Development and Long-Range Objectives for 2035 (中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标) ⁷⁰ | 2021 | <p>Calls to:</p> <ul style="list-style-type: none"> Apply genetic research for technological innovations. Establish national laboratories with a focus on AI, biotechnology, and pharmaceuticals. Reorganize the state key laboratories to form a laboratory system. Name biotechnology a strategic emerging industry. Accelerate deployment of cutting-edge technologies. |
| Personal Information Protection Law (中华人民共和国个人信息保护法) ⁷¹ | 2021 | <ul style="list-style-type: none"> Protects individuals' privacy by requiring consent to collect or process personal data, including medical and biometric information. Prevents personal data from transferring outside of China. |

| | | |
|--|-------------|--|
| <p>14th Five-Year Plan for National Informatization (“十四五”国家信息化规划)⁷²</p> | <p>2021</p> | <ul style="list-style-type: none"> ● Requires hospitals to digitize health records. ● Provides universal digital healthcare. ● Establish standardization of data. |
| <p>Guiding Opinions on Expanding Investment in Strategic Emerging Industries and Cultivating Strengthened New Growth Points and Growth Poles (关于扩大战略性新兴产业投资 培育壮大新增长点增长极的指导意见)⁷³</p> | <p>2020</p> | <p>Calls to:</p> <ul style="list-style-type: none"> ● Accelerate pace of innovation in the biotechnology industry. ● Establish biotechnology and pharmaceutical innovation centers. ● Create markets for pharmaceutical drugs and medical equipment. |
| <p>Biosecurity Law (中华人民共和国生物安全法)⁷⁴</p> | <p>2020</p> | <ul style="list-style-type: none"> ● Strengthens the management and oversight of the collection, storage, use, and external provision of Human Genetic Resources. ● Allows the government to “enjoy sovereignty” over Human Genetic Resources and biological resources. |
| <p>Regulations on the Administration of Human Genetic Resources (中华人民共和国人类遗传资源管理条例)⁷⁵</p> | <p>2019</p> | <ul style="list-style-type: none"> ● Organizations must apply for a certification to use Human Genetic Resources within China. ● Prohibits collection or preservation of Human Genetic Resources by foreign organizations and individuals. ● Foreign organizations must have a domestic sponsor to apply or use any Human Genetic Resources within China. |
| <p>Measures for the Management of Scientific Data (科学数据管理办法)⁷⁶</p> | <p>2018</p> | <ul style="list-style-type: none"> ● Requires scientific data generated in China that is considered relevant to national security issues be consolidated in newly-developed scientific data centers. |

| | | |
|--|------|---|
| Opinions on Promoting the Development of “Internet + Healthcare” (促进“互联网+医疗健康”发展的意见) ⁷⁷ | 2018 | <ul style="list-style-type: none"> ● Enhances public health services by integrating technology into hospitals and public health. ● Part of the “Internet +” initiative. |
| AI Innovation Action Plan for Institutions of Higher Education (高等学校人工智能创新行动计划) ⁷⁸ | 2018 | <p>Calls universities to:</p> <ul style="list-style-type: none"> ● Promote the deep integration of information technologies with modern biotechnology. ● Apply AI to intelligent medical care. |
| New Generation Artificial Intelligence Development Plan (新一代人工智能发展规划) ⁷⁹ | 2017 | <ul style="list-style-type: none"> ● Directs organizations to develop new models and methods of AI in fields including healthcare, precision medicine, and genetics. |
| Chinese Population Precision Medicine Research Program (中国人群精准医学研究计划) ⁸⁰ | 2016 | <ul style="list-style-type: none"> ● Funds a \$9.2 billion, 15-year program to “build up the country’s credentials in precision medicine.” ● Establishes precision medicine centers between hospitals, universities, and sequencing companies. ● Plans to sequence over 100 million human genomes to beat the same goal as the United States’ Precision Medicine Initiative. ● Establish a database of genomic data from Chinese and international populations. |
| Healthy China 2030 Plan (“健康中国2030”规划) ⁸¹ | 2016 | <p>Calls to:</p> <ul style="list-style-type: none"> ● Promote the construction of a healthy China through reform, innovation, and scientific development. ● Collect, integrate, and share data to promote big data-informed health applications. |

| | | |
|--|-------------------|--|
| National 13th 5-Year Plan for S&T Innovation (“十三五”国家科技创新规划) ⁸² | 2016 - 2020 | <p>Calls to:</p> <ul style="list-style-type: none"> ● Develop cutting-edge and general-purpose biotechnology, including gene editing and synthetic biology. ● Seize a commanding position in international biotechnology competition. |
| National 13th Five-Year Plan for the Development of Strategic Emerging Industries (“十三五”国家战略性新兴产业发展规划) ⁸³ | 2016 - 2020 | <p>Calls to:</p> <ul style="list-style-type: none"> ● Accelerate development of genomics and associated technologies for use in precision medicine, biosynthesis, and agriculture. ● Promote disruptive technological innovation in agricultural biology. ● Promote the clinical application and industrialization of artificial biology and artificial biological devices. |
| Made in China 2025 (中国制造2025) ⁸⁴ | 2015 | <p>Calls to:</p> <ul style="list-style-type: none"> ● Integrate new generation information technology and manufacturing, including in science and technology innovations like bioengineering. |
| The National Medium- and Long-Term Program for Science and Development (国家中长期科学和技术发展规划) ⁸⁵ | 2006 - 2020 | <ul style="list-style-type: none"> ● Prioritizes development of biotechnology and life sciences. ● Establishes mega-projects to meet development goals of China, including in public health and genetically-modified agriculture. ● Calls for research and development of genome sequencing and genetic structure analysis. ● Calls for breakthroughs in the fields of functional genome, proteomics, stem cells, etc. |

Source: CSET analysis. Current as of September 2023.

Authors

Caroline Schuerger is a former research fellow at the Center for Security and Emerging Technology (CSET), where Vikram Venkatram is a research analyst, and Katherine Quinn is a data scientist.

Acknowledgments

For feedback and assistance, we would like to thank Anna Puglisi, Steph Batalis, Daniel Chou, Emily Weinstein, Catherine Aiken, Owen Daniels, Ben Murphy, Shelton Fitch, Jason Ly, Margarita Konaev, and Igor Mikolic-Torreira. The authors are especially grateful to Jeff Ding and Mark Kazmierczak for their comprehensive reviews.



© 2024 by the Center for Security and Emerging Technology. This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.

To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc/4.0/>.

Document Identifier: doi: 10.51593/20220018

Endnotes

¹ Sean Mann, Carl T. Berdahl, Lawrence Baker, Federico Girosi, and Rand Health, “Applications Used in the Clinical and Public Health Response to COVID-19,” Patient-Centered Outcomes Research Institute, December 2022, <https://www.pcori.org/sites/default/files/PCORI-Emerging-Technologies-Therapeutics-Reports-AI-In-Response-COVID-19-Volume-1.pdf>.

² Marcy E Gallo, “The Bioeconomy: A Primer,” Congressional Research Service, September 22, 2022, <https://crsreports.congress.gov/product/pdf/R/R46881/4>; Market Research Future “Artificial Intelligence in Healthcare Market Size Worth USD 12.22 Billion by 2030 at 51.9% CAGR – Report by Market Research Future (MRFR),” GlobalNewswire, January 24, 2023, <https://www.globenewswire.com/en/news-release/2023/01/24/2594408/0/en/Artificial-Intelligence-in-Healthcare-Market-Size-Worth-USD-12-22-Billion-by-2030-at-51-9-CAGR-Report-by-Market-Research-Future-MRFR.html>; Michael Chui, Matthias Evers, James Manyika, Alice Zheng, and Travers Nisbet, “The Bio Revolution: Innovations transforming economies, societies, and our lives,” McKinsey & Company, May 13, 2020, <https://www.mckinsey.com/industries/life-sciences/our-insights/the-bio-revolution-innovations-transforming-economies-societies-and-our-lives>.

³ Ben Buchanan, “The AI Triad and What It Means for National Security Strategy,” Center for Security and Emerging Technology, August 2020, <https://cset.georgetown.edu/publication/the-ai-triad-and-what-it-means-for-national-security-strategy/>.

⁴ Bethany Allen-Ebrahimian, “China Makes Genetic Data a National Resource,” Axios, March 29, 2022, <https://www.axios.com/2022/03/29/china-makes-genetics-data-national-resource>; Jessie Yeung, “China’s Sitting on a Goldmine of Genetic Data – and It Doesn’t Want to Share,” CNN, August 12, 2023, <https://www.cnn.com/2023/08/11/china/china-human-genetic-resources-regulations-intl-hnk-dst/index.html>.

⁵ CSET Translation of “National 13th Five-Year Plan for the Development of Strategic Emerging Industries” (“十三五”国家战略性新兴产业发展规划), State Council (国务院), November 29, 2016, <https://cset.georgetown.edu/publication/national-13th-five-year-plan-for-the-development-of-strategic-emerging-industries/>; Infervision, “How AI And Deep Learning Are Now Used To Diagnose Cancer- Infervision,” accessed September 5, 2023, <https://global.infervision.com/news/48.html>.

⁶ Wilson Wen Bin Goh, and Limsoon Wong, “The Birth of Bio-Data Science: Trends, Expectations, and Applications,” Genomics, Proteomics & Bioinformatics, February 2020, <https://doi.org/10.1016/j.gpb.2020.01.002>.

⁷ Genome.gov, “Introduction to Genomics,” accessed September 5, 2023, <https://www.genome.gov/About-Genomics/Introduction-to-Genomics>.

⁸ “The Promise of Precision Medicine,” National Institutes of Health, February 5, 2020, <https://www.nih.gov/about-nih/what-we-do/nih-turning-discovery-into-health/promise-precision-medicine>.

⁹ “Introduction to Genomics: What’s a Genome?” National Institutes of Health National Human Genome Research Institute, accessed September 5, 2023, <https://www.genome.gov/About-Genomics/Introduction-to-Genomics>; “Artificial Intelligence, Machine Learning and Genomics,” National Institutes of Health National Human Genome Research Institute, accessed September 5, 2023, <https://www.genome.gov/about-genomics/educational-resources/fact-sheets/artificial-intelligence-machine-learning-and-genomics>; Žiga Avsec, “Predicting Gene Expression with AI,” Google DeepMind, October 4, 2021, <https://www.deepmind.com/blog/predicting-gene-expression-with-ai>.

¹⁰ Mohamad M. Kronfol, Mikhail G. Dozmorov, Rong Huang, Patricia W. Slattum, and Joseph L. McClay, “The Role of Epigenomics in Personalized Medicine,” Expert Review of Precision Medicine and Drug Development, January 17, 2017, <https://doi.org/10.1080/23808993.2017.1284557>; Lawrence B. Holder, M. Muksitul Haque, and Michael K. Skinner, “Machine Learning for Epigenetics and Future Medical Applications,” Epigenetics: Official Journal of the DNA Methylation Society June 21, 2017, <https://doi.org/10.1080/15592294.2017.1329068>.

¹¹ “What Is Proteomics?” EMBL-EBI, accessed September 5, 2023, <https://www.ebi.ac.uk/training/online/courses/proteomics-an-introduction/what-is-proteomics/>.

¹² “Medical Imaging,” U.S. Food and Drug Administration, 2018, <https://www.fda.gov/radiation-emitting-products/radiation-emitting-products-and-procedures/medical-imaging>; “More about MIDRC,” Medical Imaging and Data Resource Center, accessed September 5, 2023, <https://www.midrc.org/more-about-midrc>.

¹³ “Electronic Health Records,” Centers for Medicare and Medicaid Services, accessed September 5, 2023, <https://www.cms.gov/medicare/e-health/ehealthrecords>.

¹⁴ Sarah De Nigris, Massimo Craglia, Daniel Nepelski, Jiri Hradec, Emilio Gomez-Gonzales, Emilia Gomez Gutierrez, Miguel Vazquez-Prada Baillet, Riccardo Righi, Guiditta De Prato, Montserrat Lopez Cobo, Sofia Samoili, and Melisande Cardona, “AI Watch: AI Uptake in Health and Healthcare, 2020,” Publications Office of the European Union, 2020, ISBN 978-92-76-26936-6, doi:10.2760/948860, JRC122675; Angela Spatharou, Solveigh Hieronimus, and Jonathan Jenkins, “Transforming Healthcare with AI: The Impact on the Workforce and Organizations,” McKinsey & Company, March 10, 2020, <https://www.mckinsey.com/industries/healthcare/our-insights/transforming-healthcare-with-ai>.

¹⁵ UK Department for Science, Innovation, and Technology, “New £100 million fund to capitalise on AI’s game-changing potential in life sciences and healthcare,” October 29, 2023,

<https://www.gov.uk/government/news/new-100-million-fund-to-capitalise-on-ais-game-changing-potential-in-life-sciences-and-healthcare>; Lael Brainard, Neera Tanden, and Arati Prabhakar, “Delivering on the Promise of AI to Improve Health Outcomes,” White House, December 14, 2023, <https://www.whitehouse.gov/briefing-room/blog/2023/12/14/delivering-on-the-promise-of-ai-to-improve-health-outcomes/>; “EU invests €220 million to test AI solutions for healthcare, food, industry, and everyday life,” European Commission, June 29, 2023, <https://digital-strategy.ec.europa.eu/en/news/eu-invests-eu220-million-test-ai-solutions-healthcare-food-industry-and-everyday-life>.

¹⁶ The White House, “Executive Order on Advancing Biotechnology and Biomanufacturing Innovation for a Sustainable, Safe, and Secure American Bioeconomy,” September 12, 2022, <https://www.whitehouse.gov/briefing-room/presidential-actions/2022/09/12/executive-order-on-advancing-biotechnology-and-biomanufacturing-innovation-for-a-sustainable-safe-and-secure-american-bioeconomy/>.

¹⁷ World Bank Open Data, “Physicians (per 1,000 People.),” accessed September 5, 2023, <https://data.worldbank.org/indicator/SH.MED.PHYS.ZS?locations=CN-US-AU-BR-TR>; Sarah O’Meara, “China’s Data-Driven Dream to Overhaul Health Care,” *Nature* 598 (7879): S1–3, 2021, <https://www.nature.com/articles/d41586-021-02694-1>.

¹⁸ World Food Programme, “China,” accessed September 5, 2023, <https://www.wfp.org/countries/china>; Ethan Wang and Bernard Orr, “China to Promote Vaccine Development, New Medicines,” Reuters, March 5, 2023, <https://www.reuters.com/world/china/china-promote-vaccine-development-new-medicines-2023-03-05/>; Dominique Patton, “Exclusive: China Rolls out GMO Corn Planting, Starts Small,” Reuters, February 17, 2023, <https://www.reuters.com/markets/commodities/china-rolls-out-gmo-corn-planting-starts-small-2023-02-16/>; Wang, Long-De, Yu Wang, Gong-Huan Yang, Jia-Qi Ma, Li-Ping Wang, and Xiao-Peng Qi, “China Information System for Disease Control and Prevention (CISDCP),” National Bureau of Asian Research, accessed September 5, 2023, <https://www.nbr.org/wp-content/uploads/pdfs/programs/CISDCP.pdf>.

¹⁹ “Yidu Cloud,” accessed September 5, 2023, <https://yiduccloud.com.cn/en/solution.html>.

²⁰ Lulu Yilun Chen, “Tencent-Backed WeDoctor Plans IPO Filing by End of April” Bloomberg, January 8, 2023, <https://www.bloomberg.com/news/articles/2023-01-09/tencent-backed-wedoctor-plans-ipo-filing-by-end-of-april>.

²¹ Through Healthy China 2030, the government has reportedly invested over \$9 billion USD for the Chinese Population Precision Medicine Research Program (中国人群精准医学研究计划) to utilize the wealth of accumulated health biodata, and to “realize data collection, integration, sharing, and business of public health.” International Telecommunication Union (ITU) translation of “The National Medium- and

Long-Term Program for Science and Technology Development (2006-2020): An Outline” (国家中长期科学和技术发展规划纲要（2006—2020年）), State Council, 2006, https://www.itu.int/en/ITU-D/Cybersecurity/Documents/National_Strategies_Repository/China_2006.pdf; “The Chinese Communist Party Central Committee and the State Council Issue *Outline of the ‘Healthy China 2030’ Plan*” (中共中央 国务院印发《“健康中国2030”规划纲要》), Xinhua News Agency, October 25, 2016, <https://perma.cc/8576-U8HX>.

²² David Cyranoski, “China Embraces Precision Medicine on a Massive Scale.” *Nature* 529 (7584): 9–10, January 6, 2016, <https://doi.org/10.1038/529009a>.

²³ “All of Us Research Program Overview,” National Institute of Health All of Us Research Program, June 22, 2020, <https://allofus.nih.gov/about/program-overview>.

²⁴ “Translation: 14th Five-Year Plan for National Informatization—Dec. 2021,” DigiChina, <https://digichina.stanford.edu/work/translation-14th-five-year-plan-for-national-informatization-dec-2021/>.

²⁵ “The Chinese Communist Party Central Committee and the State Council Issue *Outline of the ‘Healthy China 2030’ Plan*” (中共中央 国务院印发《“健康中国2030”规划纲要》), Xinhua News Agency, October 25, 2016, <https://perma.cc/8576-U8HX>.

²⁶ China Law Translate translation of “Biosecurity Law of the P.R.C.” (中华人民共和国生物安全法), National People’s Congress Standing Committee (国家人民代表大会常务委员会), October 17, 2020, <https://perma.cc/VM3X-DW9X>; CSET translation of “Measures for the Management of Scientific Data” (科学数据管理办法), General Office of the State Council (国务院办公厅), April 2, 2018, <https://cset.georgetown.edu/publication/china-scientific-data-management-measures/>.

²⁷ “Privacy in Genomics,” National Institutes of Health National Human Genome Research Institute, accessed September 5, 2023, <https://www.genome.gov/about-genomics/policy-issues/Privacy>.

²⁸ Emile Dirks and James Leibold, “Genomic Surveillance,” Australian Strategic Policy, June 17, 2020, <https://www.aspi.org.au/report/genomic-surveillance>; Sophie Richardson, “China: Minority Region Collects DNA from Millions,” Human Rights Watch, December 13, 2017, <https://www.hrw.org/news/2017/12/13/china-minority-region-collects-dna-millions>.

²⁹ United States Department of the Treasury, Office of Foreign Assets, “Treasury Sanctions Chinese Entity and Officials Pursuant to Global Magnitsky Human Rights Executive Order,” July 30, 2020, <https://home.treasury.gov/news/press-releases/sm1073>; Sophie Richardson, “China: Minority Region Collects DNA from Millions,” Human Rights Watch, December 13, 2017, <https://www.hrw.org/news/2017/12/13/china-minority-region-collects-dna-millions>; Emile Dirks and James Leibold, “Genomic Surveillance,” Australian Strategic Policy, June 17, 2020,

<https://www.aspi.org.au/report/genomic-surveillance>; Emile Dirks, “Mass DNA Collection in the Tibet Autonomous Region from 2016–2022,” The Citizen Lab, September 13, 2022, <https://citizenlab.ca/2022/09/mass-dna-collection-in-the-tibet-autonomous-region/>.

³⁰ Yi Li, Wenting Zhao, Dan Li, Xianming Tao, Ziyi Xiong, Jing Liu, Wei Zhang, et al., “RETRACTED ARTICLE: EDAR, LYPLAL1, PRDM16, PAX3, DKK1, TNFSF12, CACNA2D3, and SUPT3H Gene Variants Influence Facial Morphology in a Eurasian Population,” *Human Genetics* 138 (6): 681–89 April 25, 2019, <https://doi.org/10.1007/s00439-019-02023-7>; Xiaoxi Jing, Yanan Sun, Wenting Zhao, Xingjian Gao, Mi Ma, Fan Liu, and Caixia Li, “RETRACTED ARTICLE: Predicting Adult Height from DNA Variants in a European-Asian Admixed Population,” *International Journal of Legal Medicine* 133 (6): 1667–79, April 12, 2019, <https://doi.org/10.1007/s00414-019-02039-8>.

³¹ “China National GeneBank (CNGB)- About” BGI Group Official Website, accessed September 5, 2023, <https://en.genomics.cn/hwzp>.

³² MGI Tech, “MGI - Genetic Sequencing and Multi-Omics Solutions for Life Science and Healthcare Research,” accessed September 5, 2023, <https://en.mgi-tech.com/>.

³³ “Complete Genomics, BGI Agree to \$117.6M Merger,” *GenomeWeb*, September 17, 2012, <https://www.genomeweb.com/clinical-sequencing/complete-genomics-bgi-agree-1176m-merger>.

³⁴ Current CLIA Identification number is 99D2135851, and expires on 09/07/2025; Reuters. “BRIEF-BGI Genomics Unit Receives CLIA Certificate,” September 11, 2017, <https://www.reuters.com/article/brief-bgi-genomics-unit-receives-clia-ce-idUKL4N1LS20K>; Centers for Medicare and Medicaid Services, “Clinical Laboratory Improvement Amendments (CLIA),” accessed September 5, 2023, <https://www.cms.gov/regulations-and-guidance/legislation/clia>.

³⁵ Kirsty Needham, “Exclusive: China Gene Firm Providing Worldwide COVID Tests Worked with Chinese Military,” Reuters, January 30, 2021, <https://www.reuters.com/article/us-china-genomics-military-exclusive-idUSKBN29Z0HA>; Kirsty Needham and Clare Baldwin, “China’s Gene Giant Harvests Data from Millions of Pregnant Women,” Reuters, July 7, 2021, <https://www.reuters.com/investigates/special-report/health-china-bgi-dna/>.

³⁶ Kirsty Needham, “Exclusive: China Gene Firm Providing Worldwide COVID Tests Worked with Chinese Military.”; Kirsty Needham and Clare Baldwin, “China’s Gene Giant Harvests Data from Millions of Pregnant Women.”

³⁷ “Translation: 14th Five-Year Plan for National Informatization—Dec. 2021,” DigiChina, <https://digichina.stanford.edu/work/translation-14th-five-year-plan-for-national-informatization-dec-2021/>.

- ³⁸ Chinese Academy of Sciences, “Big Gene Bank to Anchor Precision Medicine,” Chinese Academy of Sciences, December 9, 2015, https://english.cas.cn/newsroom/archive/china_archive/cn2015/201512/t20151209_157214.shtml; Chinese National Genebank, National Genomics Data Center, “About - National Genomics Data Center (CNGB - NGDC),” accessed September 5, 2023, <https://ngdc.cnbc.ac.cn/about>.
- ³⁹ “GenBank Overview,” National Institutes of Health, National Library of Medicine, National Center for Biotechnology Information, accessed September 5, 2023, <https://www.ncbi.nlm.nih.gov/genbank/>; UK Biobank, “About UK Biobank,” accessed September 5, 2023, <https://www.ukbiobank.ac.uk/>.
- ⁴⁰ Jessie Yeung, “China’s Sitting on a Goldmine of Genetic Data -- and It Doesn’t Want to Share,” CNN, August 12, 2023, <https://www.cnn.com/2023/08/11/china/china-human-genetic-resources-regulations-intl-hnk-dst/index.html>.
- ⁴¹ “GenBank Overview,” National Institutes of Health, National Library of Medicine, National Center for Biotechnology Information, accessed January 19, 2024, <https://www.ncbi.nlm.nih.gov/genbank/>; RCSB Protein Data Bank, “About RCSB PDB: Enabling Breakthroughs in Scientific and Biomedical Research and Education,” accessed January 19, 2024, <https://www.rcsb.org/pages/about-us/index>.
- ⁴² Center for Open Data Enterprise, “Sharing and Utilizing Health Data for AI Applications: Roundtable Report,” 2019, <https://www.hhs.gov/sites/default/files/sharing-and-utilizing-health-data-for-ai-applications.pdf>.
- ⁴³ Eric Wu, Kevin Wu, Roxana Daneshjou, David Ouyang, Daniel E. Ho, and James Zou, “How Medical AI Devices Are Evaluated: Limitations and Recommendations from an Analysis of FDA Approvals,” *Nature Medicine* 27 (4): 582–84, 2017, April 5, 2021, <https://doi.org/10.1038/s41591-021-01312-x>; Shana Lynch, “The Geographic Bias in Medical AI Tools,” Stanford University Human-Centered Artificial Intelligence, September 21, 2020, <https://hai.stanford.edu/news/geographic-bias-medical-ai-tools>; Center for Open Data Enterprise, “Sharing and Utilizing Health Data for AI Applications: Roundtable Report.”
- ⁴⁴ “Translation: Outbound Data Transfer Security Assessment Measures – Effective Sept. 1, 2022,” DigiChina, <https://digichina.stanford.edu/work/translation-outbound-data-transfer-security-assessment-measures-effective-sept-1-2022/>.
- ⁴⁵ Jun-Yu Xu, Chunchao Zhang, Xiang Wang, Linhui Zhai, Yiming Ma, Yousheng Mao, Kun Qian, et al., “Integrative Proteomic Characterization of Human Lung Adenocarcinoma,” *Cell* 182 (1): 245-261.e17, July 19, 2020, <https://doi.org/10.1016/j.cell.2020.05.043>.

⁴⁶ Office for Civil Rights (OCR), “Your Rights under HIPAA.” U.S. Department of Health and Human Services, May 7, 2008, <https://www.hhs.gov/hipaa/for-individuals/guidance-materials-for-consumers/index.html>.

⁴⁷ CSET translation of “National 13th Five-Year Plan for the Development of Strategic Emerging Industries” (“十三五”国家战略性新兴产业发展规划), State Council (国务院), November 29, 2016, <https://cset.georgetown.edu/publication/national-13th-five-year-plan-for-the-development-of-strategic-emerging-industries/>.

⁴⁸ CSET translation of “Notice of the Ministry of Education on the Publication of the *AI Innovation Action Plan for Institutions of Higher Education*” (教育部关于印发《高等学校人工智能创新行动计划》的通知), April 2, 2018, <https://cset.georgetown.edu/research/ai-innovation-action-plan-for-institutions-of-higher-education>.

⁴⁹ Anna Puglisi and Daniel Chou, “China’s Industrial Clusters: Building AI-Driven Bio-Discovery Capacity,” Center for Security and Emerging Technology, June 2022, <https://cset.georgetown.edu/publication/chinas-industrial-clusters/>; Emily S. Weinstein, Channing Lee, Ryan Fedasiuk, and Anna Puglisi, “China’s State Key Laboratory System: A View into China’s Innovation System,” Center for Security and Emerging Technology, June 2022, <https://cset.georgetown.edu/publication/chinas-state-key-laboratory-system/>.

⁵⁰ Julia Karow, “BGI to Receive \$1.5B in ‘collaborative Funds’ over 10 Years from China Development Bank,” GenomeWeb, January 12, 2010, <https://www.genomeweb.com/sequencing/bgi-receive-15b-collaborative-funds-over-10-years-china-development-bank>.

⁵¹ Anna Puglisi and Daniel Chou, “China’s Industrial Clusters: Building AI-Driven Bio-Discovery Capacity.”

⁵²The National Medical Products Administration (NMPA; 国家药品监督管理局) is the agency in China that is responsible for approving and certifying medical devices and pharmaceutical drugs.

⁵³ National Medical Products Association, “Key Points of Deep Learning Decision-Assisting Medical Device Software Review” Archived September 11, 2022, <https://perma.cc/3VVC-7J3B>.

⁵⁴ CSET merged corpus of scholarly literature including [Web of Science](#), OpenAlex, Semantic Scholar, [The Lens](#), arXiv, and Papers With Code. Certain data included herein are derived from Clarivate Web of Science. © Copyright Clarivate 2023. All rights reserved. The majority of publications identified from our keyword search were English-language, though we did include Chinese-language keywords and the search did find some relevant papers in Chinese and other languages.

⁵⁵ A publication is counted for a country if at least one of its authors is affiliated with a research institution in that country. For example, a paper with one author affiliated with a Chinese institution and one author affiliated with a U.S. institution is added as one to each country's paper count. For more discussion on this counting method, see Jacob Feldgoise, Catherine Aiken, Emily S. Weinstein, and Zachary Arnold, "Studying Tech Competition through Research Output: Some CSET Best Practices," Center for Security and Emerging Technology, April 2023, <https://cset.georgetown.edu/article/studying-tech-competition-through-research-output-some-cset-best-practices/>.

⁵⁶ Eleanor Olcott, "China Sets the Pace in Adoption of AI in Healthcare Technology," *Financial Times*, January 31, 2022, <https://www.ft.com/content/c1fe6fbf-8a87-4328-9e75-816009a07a59>.

⁵⁷ Nature Advertisement, "Spearheading a New Age of Medical Intelligence," Nature, accessed September 5, 2023, <https://www.nature.com/articles/d42473-021-00188-2>. Yidu Cloud, "Medical Big Data Intelligent Solutions," accessed September 5, 2023, <https://yiduccloud.com.cn/en/solution.html>; Yidu Cloud "Yidu Tech Announces Annual Results for Fiscal Year 2022," accessed September 5, 2023, <https://www.yidutechgroup.com/en/news-detail.html?id=143>; Yidu Cloud, "Yidu Tech Leads China's Healthcare Intelligence Market," accessed September 5, 2023, <https://www.yidutechgroup.com/en/news-detail.html?id=142>.

⁵⁸ Microsoft, "AI and Preventative Healthcare: Diagnosis in the Blink of an Eye," Microsoft Stories Asia, September 17, 2018, <https://news.microsoft.com/apac/features/ai-and-preventative-healthcare-diagnosis-in-the-blink-of-an-eye/>; Duoru Lin, Jianhao Xiong, Congxin Liu, Lanqin Zhao, Zhongwen Li, Shanshan Yu, Xiaohang Wu, et al., "Application of Comprehensive Artificial Intelligence Retinal Expert (CARE) System: A National Real-World Evidence Study." *The Lancet. Digital Health* 3 (8): e486–95, August 2021, [https://doi.org/10.1016/s2589-7500\(21\)00086-8](https://doi.org/10.1016/s2589-7500(21)00086-8); National Medical Products Administration, accessed September 2023, <https://www.nmpa.gov.cn/xxgk/ggtg/ylqxpzhzhcchpgg/20200916150323159.html>.

⁵⁹ InferVision, "InferVision Receives FDA Clearance for the InferRead Lung CT.AI Product," accessed January 24, 2024, <https://us.infervision.com/news/68.html>.

⁶⁰ InferVision, "How AI And Deep Learning Are Now Used To Diagnose Cancer-InferVision," accessed September 5, 2023, <https://global.infervision.com/news/48.html>.

⁶¹ InferVision, "How AI And Deep Learning Are Now Used To Diagnose Cancer-InferVision"; Emily S. Weinstein, "China's Use of AI in its COVID-19 Response," Center for Security and Emerging Technology, August 2020, <https://cset.georgetown.edu/publication/chinas-use-of-ai-in-its-covid-19-response/>.

⁶² Baidu, “How Baidu Is Bringing AI to the Fight against Coronavirus,” *MIT Technology Review*, March 11, 2020, <https://www.technologyreview.com/2020/03/11/905366/how-baidu-is-bringing-ai-to-the-fight-against-coronavirus/>.

⁶³ Baidu, “Baidu’s Deep-Learning Platform Fuels the Rise of Industrial AI,” *MIT Technology Review*, June 22, 2020, <https://www.technologyreview.com/2020/06/22/1004251/baidus-deep-learning-platform-fuels-the-rise-of-industrial-ai/>; Baidu, “How Baidu Is Bringing AI to the Fight against Coronavirus”; Liang Huang, He Zhang, Dezhong Deng, Kai Zhao, Kaibo Liu, David A. Hendrix, and David H. Mathews, “LinearFold: Linear-Time Approximate RNA Folding by 5’-to-3’ Dynamic Programming and Beam Search,” *Bioinformatics (Oxford, England)* 35 (14): i295–304 July 2019, <https://doi.org/10.1093/bioinformatics/btz375>.

⁶⁴ NASDAQ, “How Alibaba Is Using Artificial Intelligence In Healthcare,” July 13, 2017, <https://www.nasdaq.com/articles/how-alibaba-using-artificial-intelligence-healthcare-2017-07-13>; Alibaba Cloud “Medical Brain: Data Intelligence for Healthcare & Lives Saving - Alibaba Cloud,” accessed September 5, 2023, <https://www.alibabacloud.com/solutions/intelligence-brain/medical?spm=a3c0i.14008940.5178217040.2.21397bffFo01f>.

⁶⁵ Tencent, “Jarvis Lab,” accessed September 5, 2023, <https://jarvislab.tencent.com/index-en.html>; Lulu Yilun Chen, “Tencent-Backed WeDoctor Plans IPO Filing by End of April” Bloomberg, January 8, 2023, <https://www.bloomberg.com/news/articles/2023-01-09/tencent-backed-wedoctor-plans-ipo-filing-by-end-of-april?leadSource=uverify%20wall#xj4y7vzkg>; Novartis, “AI Nurse Set to Transform Standards of Care for Heart Failure Patients in China,” April 30, 2020, <https://www.novartis.com/news/ai-nurse-set-transform-standards-care-heart-failure-patients-china>; Tencent, “Tencent Announces AIMIS Medical Image Cloud and AIMIS Open Lab Help Medical Data Management and Accelerate Incubation of Medical AI Application,” PR Newswire, October 20, 2020, <https://www.prnewswire.com/news-releases/tencent-announces-aimis-medical-image-cloud-and-aimis-open-lab-help-medical-data-management-and-accelerate-incubation-of-medical-ai-application-301155456.html>; Benjamin Larsen, “Drafting China’s National AI Team for Governance,” DigiChina, November 18, 2019, <https://digichina.stanford.edu/work/drafting-chinas-national-ai-team-for-governance/>.

⁶⁶ Yuchen Li, “The Misfires: How BAT All Stumbled in Medical AI,” Translated by Wenmiao Liu, August 6 2021, *Leiphone*, https://docs.google.com/document/d/19PykMYFsVf1tLrkxemv9oc2DvAmkBLTBq7vS_BWSd8c/edit.

⁶⁷ U.S. Food and Drug Administration, “Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices,” October 5, 2022, <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>.

⁶⁸ China National Medical Products Administration, “Guidelines for the Registration and Examination of AI Medical Devices,” <https://perma.cc/U8T5-9N4U>.

⁶⁹ “Translation: Outbound Data Transfer Security Assessment Measures – Effective Sept. 1, 2022,” DigiChina, <https://digichina.stanford.edu/work/translation-outbound-data-transfer-security-assessment-measures-effective-sept-1-2022/>.

⁷⁰ CSET translation of “Outline of the People’s Republic of China 14th Five-Year Plan for National Economic and Social Development and Long-Range Objectives for 2035” (中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲), Xinhua News Agency (新华社), March 12, 2021, <https://cset.georgetown.edu/publication/china-14th-five-year-plan>.

⁷¹ “Translation: Personal Information Protection Law of the People’s Republic of China—Effective Nov. 1, 2021,” DigiChina blog, <https://digichina.stanford.edu/work/translation-personal-information-protection-law-of-the-peoples-republic-of-china-effective-nov-1-2021>.

⁷² “Translation: 14th Five-Year Plan for National Informatization—Dec. 2021,” DigiChina, <https://digichina.stanford.edu/work/translation-14th-five-year-plan-for-national-informatization-dec-2021/>.

⁷³ CSET-DigiChina translation of “Guiding Opinions on Expanding Investment in Strategic Emerging Industries and Cultivating Strengthened New Growth Points and Growth Poles” (关于扩大战略性新兴产业投资 培育壮大新增长点增长极的指导意见), National Development and Reform Commission (国家发展和改革委员会), September 8, 2020, <https://cset.georgetown.edu/publication/new-chinese-ambitions-for-strategic-emerging-industries-translated>.

⁷⁴ China Law Translate translation of “Biosecurity Law of the P.R.C.” (中华人民共和国生物安全法), National People’s Congress Standing Committee (国家人民代表大会常务委员会), October 17, 2020, <https://perma.cc/VM3X-DW9X>.

⁷⁵ Ministry of Science and Technology (MOST; 科学技术部), “Regulations on the Administration of Human Genetic Resources of the People’s Republic of China” (中华人民共和国人类遗传资源管理条例), May 28, 2019, <https://perma.cc/WZ76-K3R6>.

⁷⁶ CSET translation of “Measures for the Management of Scientific Data” (科学数据管理办法), General Office of the State Council (国务院办公厅), April 2, 2018, <https://cset.georgetown.edu/publication/china-scientific-data-management-measures/>.

⁷⁷ “Opinions of the General Office of the State Council on Promoting the Development of ‘Internet + Healthcare’” (国务院办公厅关于促进“互联网+医疗健康”发展的意见), April 25, 2018, <https://perma.cc/47TT-YGAG>.

⁷⁸ CSET translation of “Notice of the Ministry of Education on the Publication of the *AI Innovation Action Plan for Institutions of Higher Education*” (教育部关于印发《高等学校人工智能创新行动计划》的通知), April 2, 2018, <https://cset.georgetown.edu/research/ai-innovation-action-plan-for-institutions-of-higher-education>.

⁷⁹ “Full Translation: China’s ‘New Generation Artificial Intelligence Development Plan’ (2017),” DigiChina blog, <https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/>.

⁸⁰ Brian Wang, “China’s \$9.2 Billion Precision Medicine Initiative Could See about 100 Million Whole Human Genomes Sequenced by 2030 and More If Sequencing Costs Drop,” Nextbigfuture.com, June 7, 2016; and David Cyranoski, “China Embraces Precision Medicine on a Massive Scale,” *Nature* 529 (2016): 9–10.

⁸¹ “The Chinese Communist Party Central Committee and the State Council Issue *Outline of the ‘Healthy China 2030’ Plan*” (中共中央 国务院印发《“健康中国2030”规划纲要》), Xinhua News Agency, October 25, 2016, <https://perma.cc/8576-U8HX>.

⁸² CSET translation of “State Council Notice on the Publication of the National 13th Five-Year Plan for S&T Innovation” (国务院关于印发“十三五”国家科技创新规划的通知), Central People’s Government of the People’s Republic of China, July 28, 2016, <https://cset.georgetown.edu/research/state-council-notice-on-the-publication-of-the-national-13th-five-year-plan-for-st-innovation>.

⁸³ CSET translation of “National 13th Five-Year Plan for the Development of Strategic Emerging Industries” (国务院关于印发“十三五”国家战略性新兴产业发展规划的通知), Central People’s Government of the People’s Republic of China, November 29, 2016, <https://cset.georgetown.edu/research/national-13th-five-year-plan-for-the-development-of-strategic-emerging-industries/>.

⁸⁴ CSET translation of “Notice of the State Council on the Publication of Made in China 2025” (国务院关于印发《中国制造2025》的通知), Central People’s Government of the People’s Republic of China, May 8, 2015, <https://cset.georgetown.edu/publication/notice-of-the-state-council-on-the-publication-of-made-in-china-2025>.

⁸⁵ International Telecommunication Union (ITU) translation of “The National Medium- and Long-Term Program for Science and Technology Development (2006-2020): An Outline” (国家中长期科学和技术发展规划纲要（2006—2020年）), State Council, 2006, https://www.itu.int/en/ITU-D/Cybersecurity/Documents/National_Strategies_Repository/China_2006.pdf.