

Study on the Protection of Human Genetic Information in the Era of Big Biological Data

Meixuan Xu^{1,*}

¹ Law School, Huazhong University of Science and Technology, Wuhan, Hubei, China

*Corresponding author. Email: meixuanxu@hust.edu.cn

ABSTRACT

The overlapping development of life sciences and information sciences has promoted the arrival of biological big data era. As one of the fast-growing big data types, biological big data has been profoundly affecting biomedicine and promoting the popularization and application of "precision medicine". However, biological information data cannot exist independently of the basic resource of Biobank. Moreover, the Biobank, which stores human genetic information, is very special. If the human genetic information cannot be strictly protected, it will inevitably cause a shake of the person's subjective status. Therefore, it is very necessary to rationally examine the harm caused by the unreasonable disclosure of current genetic information, and adopt effective countermeasures for the realization the benign operation of the Biobank and enhancement of the welfare of our people.

Keywords: *Big biological data, Biobank, Genetic information, Genetic discrimination*

1. INTRODUCTION

The coming of big biological data era has benefited from the emergence and development of bioinformatics. "Biological big data" is the result of the concept of "big data" in recent years and has become a scientific term often used by biologists. The reason why it is called "biological big data" not only refers to the huge volume of data, but continuous flow, wide sharing of data and the creation of a series of value that benefits the whole society.

Big biological data has become so important that it is inseparable from the deepening of the concept of "precision medicine". According to Zhan Qimin, an academician of the Chinese Academy of Engineering, precision medicine should run through the entire process of disease, including disease risk prediction, diagnosis and classification, drug application, efficacy evaluation, and prognosis prediction. The core of precision medicine is "precision", and to be precise, biological information is the key element. Precision medicine is an emerging medical paradigm. Its future success or failure depends on the completeness of the construction of large-scale Biobanks. With the rapid popularization of the concept of precision medicine, major countries in the world, including China, have established larger strategic goals for Biobanks.

From a macro perspective, the establishment of Biobanks is an important infrastructure for scientific research in developing countries such as China. The good operation of this infrastructure can provide convenient conditions for subsequent research and development, transformation of results, and commercial applications. In addition, the accelerated development of Biobanks will also provide resource guarantees for the enhancement of national public service functions. Biobank is an important basic platform for industrial and commercial applications driven by academic research and industrial development. However, all relevant links in the construction process of Biobanks involve genetic information, which is the core element of Biobanks.

2. THE REGULATIONS ON HUMAN GENETIC INFORMATION IN CIVIL CODE OF THE PEOPLE'S REPUBLIC OF CHINA

In recent years, China's laws have made a series of regulations on the protection of human genetic information. Article 1034 of the Personality Rights of the Civil Code stipulates that a natural person's personal information is protected by law, and the relevant definition of personal information is clarified. Personal biometric information with a high degree of personality attributes and physical orientation belongs to the

category of personal information and is protected by national laws. Facial features, fingerprints, iris, voice, genes, gait, handwriting and other information that can identify the physiological and behavioral characteristics of natural persons can all be called personal biometric information. As typical biometric information, that is, human genetic information stored in "DNA" that can completely reflect biological characteristics, it should be protected.

Article 1009 of the Personality Rights of the Civil Code created the private law norms for the protection of human genes. A medical and scientific research activity related to human genes, embryos, or the like, shall be done in accordance with the relevant provisions of laws, administrative regulations, and the regulations of the State, and shall not endanger human health, offend ethics and morals, or harm public interests. It provides a set of verification standards for the legitimacy of medical and scientific research activity related to human genes, embryos. In addition, we must resolutely resist violations of personal freedom and human dignity. In addition to the Civil Code, Amendment (11) to the Criminal Law also added articles on punishment for crimes related to human genetic resources and biosecurity. The protection of human gene rights in our country is gradually strengthened, and the governance and supervision of biomedical research activities and Biobanks will also receive more and more attention from the country in the future. [1]

3. THE SIGNIFICANCE OF PROTECTING HUMAN GENETIC INFORMATION

3.1. Resolve the privacy dilemma brought by the development of genomics

First of all, the increasing accumulation of human genetic information benefits from the continuous major breakthroughs in genetics and genomics research. However, it cannot be ignored that the rapid development of both has posed a great challenge to the protection of human genetic information and personal privacy. With the continuous iteration of emerging genetic technologies, genomics has been widely applied in the field of medicine and biotechnology, enabling clinical medical staff and biomedical researchers to obtain a large number of genetic information from large-scale patients and subjects to carry out more in-depth medical research activities. However, if the acquisition and sharing of human genetic information is not reasonably restricted, it will pose a great threat to personal privacy. Privacy is a secret that a natural person is unwilling to be known by others (people outside a certain range) in their personal lives. However, the right of privacy is a basic personality right that cannot be illegally deprived and others must

bear corresponding legal obligations to ensure the smooth realization of personal privacy.

Therefore, human genetic information itself has both intrinsic and instrumental values, and protecting human genetic information has important practical significance. Moreover, within the existing legal framework, the protection of rights related to human genetic information must be based on respect for human rights.

3.2. Correction of eugenics that ignores the intrinsic value of human beings

Secondly, although massive amounts of human genetic information can bring major advances in the diagnosis and treatment of diseases, it also raises concerns about eugenics and misuse of genetic information. Eugenics is one of the manifestations of genetic determinism, insisting that there is a natural connection between genetics and human behaviors and diseases. It argues that knowing a person's genetic sequence is equivalent to predicting someone's future.

In history, the eugenics movement in the 19th century is considered the precursor of genetic discrimination. [2] The British geneticist Francis Galton believes that eugenics is a science of optimizing race and an application form of genetics. Its goal is to allow high-quality blood or races to multiply and eliminate inferior ones. Genetic determinism is essentially a mechanistic view, which will only lead to social stratification based on genetic composition. This theory completely negates the inherent value of man. Equality and freedom are an important tenet of a civilized society, but the idea of eugenics has made the idea that all men are created equal is ruthlessly trampled on.

In short, when human beings are involved in the era of big biological data, the subjective status of human beings is constantly threatened by objectification and may face human rights challenges at any time. Therefore, it is necessary to inject the value concept of "Human beings are ends, not means", face up to the actual risks faced by human genetic information, and take seriously the governance problems of Biobanks in the development process.

4. THE TRIPLE RISKS FACED BY HUMAN GENETIC INFORMATION

4.1. Individual risk

Anti-genetic discrimination is one of the most important aspects of protection of human genetic information is. For example, *Health Insurance Portability and Accountability Act (HIPAA)*, *Genetic Information Nondiscrimination Act of 2008 (GINA)*, *Privacy Act* and other series of acts promulgated by the United States are all based on typical scenarios where

genetic discrimination may exist to formulate corresponding anti-discrimination countermeasures and protect privacy. Specifically, the typical areas where genetic discrimination may exist, such as the field of entering a higher school, are prominently manifested in the realization of the citizen's right to education. Another example is the field of employment, which is prominently manifested as the issue of employment and labor relations. Finally, in the field of insurance, it is prominently manifested in the insurance issues of commercial insurance, such as threshold of access and premium standards, etc. Genetic discrimination may even appear in areas such as visa permits and mortgage loans.

First, the field of education. Because genetic information provides key information about personal characteristics, if disclosed or abused, it may cause serious harm and cause potential discrimination in education. For example, in the field of education, children may be refused admission to education based on the assumed correlation between genetic information or genetic variation and cognitive function. On the one hand, the right to education is a fundamental right of Chinese citizens, which is confirmed and guaranteed by the Constitution. It is both a right and an obligation. On the other hand, the right to education is also a basic human right. Education plays an important role in the survival and development of citizens in the future. If they are treated unfairly in the field of education, it will have many adverse effects on the shaping of children's personality.

Second, the field of employment. Generally, when conducting biomedical research, one will consider whether the research involves disadvantaged or minority groups, gender discrimination, age discrimination, or identity discrimination. Before the establishment of a large-scale Biobank, it may only be a small-scale research or a simple survey conducted by a few people, and the collected information and data were summarized, classified, and studied, so the scope of influence was extremely limited. However, when large-scale comparative studies using Biobanks continue to advance, the in-depth links between Biobanks and social groups will cause many ethical and legal issues.

For example, when it comes to future employment, some large enterprises may force job-seekers to undergo genetic testing, or even take genetic testing as a prerequisite for employment and screen job seekers, or construct a large-scale research information database of enterprises to detect employees with certain genetic tendencies, such as detecting employees with depression tendencies in the company, and use this as a reason for dismissing the employee.

Third, the insurance field. The possible insurance discrimination caused by human genetic information has become one of the main concerns of the public. Most European countries have complete social welfare systems

and have established a relatively complete medical security system to provide medical care to all citizens, regardless of their employment status. The main source of insurance money is national taxes, so the level of medical insurance is relatively high. Take the United Kingdom as an example, the medical insurance model mainly adopts the National Health Service, or the national medical insurance system model. Due to the difference in social background, it is in sharp contrast with the United States.

China has basically established a medical security system covering the entire population. The existing medical security system is more similar to the social insurance model in Germany. The source of insurance funds is mainly paid by employers and workers in a certain proportion. At the same time, the government provides appropriate subsidies, so the government only assumes part of the social responsibility. There is still a gap between China and the United Kingdom and other countries that pursue a national medical insurance model. Our country can only solve the basic survival problems of the people, but cannot cover the main risks of the family, such as serious illnesses and accidents. In addition, there are certain restrictions on the amount of reimbursement. Therefore, ordinary families in China will still purchase certain commercial medical insurance as a reasonable supplement to social medical insurance to avoid family risks comprehensively.

In short, although genetic discrimination in the insurance field has not yet broken out on a large scale in our country, based on the irreversible harm of genetic discrimination to human dignity, our country must also take strict precautions against possible genetic discrimination in the medical insurance and life insurance fields.

4.2. Group risk

For example, a researcher or research institution intends to investigate the drinking habits of Mongolian people, which seems to be only a public health investigation. However, if a negative research conclusion reaches, related papers are published, and the research results are disclosed to the public. It is likely that the Mongolian people will be labeled with some kind of label. For example, the public defines the Mongolian people as a "Nation of alcoholic", which has caused a serious stigma problem.

Although the human genetic information provider's informed consent was obtained and the relevant informed consent documents were signed, the informed consent procedure set up by the genetic information had to be differentiated from other types of information due to the family genetic characteristics of genetic information. Furthermore, researchers or institutions need explain their the identity of the special ethnic group with the

human genetic information provider in advance, and ask whether they are willing to disclose such an identity, and what negative effects and possible risks may be brought about. These matters must be informed in detail.

Take Taiwan as an example. The ethnic group in Taiwan is not particularly large in comparison. Whenever a research institution receives biological samples or genetic information of indigenous peoples, it cannot refuse the indigenous people to participate, because this involves fair participation and ethical issues. In addition, the indigenous people will be reminded to be protected by the Basic Law of the Indigenous Peoples in the process of participating, and they will be informed that they cannot use their biological samples for biomedical research without the explicit consent of the tribe. However, research institutions can store biological samples first and wait for the tribe's consent before they can be used for further research. [3]

4.3. Country risk

The ultimate goal of establishing Biobank is to end the disease. However, it is likely to lead to fierce competition between countries, just like the "digital divide" that exists between developing and developed countries in the field of information technology, a "gene divide" may also occur in the future.

When an emerging technology is born, and it undergoes more complex ethical and legal regulations, it will increase the cost of the technology. The reason is that a technology relies on a large amount of intellectual property rights to protect a series of intellectual achievements produced by the technology. When the technological achievements are successfully transformed, commercialized and re-entered into the market, the price set may be too high for ordinary people to afford. As a result, only countries with advanced genetic technology and complete Biobanks can obtain effective drugs.

New medical drugs play an irreplaceable role in the treatment of difficult diseases. Generally speaking, developing countries are rich in biological resources and often serve as providers of biological resources, but at the same time they are often the places where AIDS, malaria and various plagues spread. When these countries are fighting diseases, they have to pay high prices for drugs due to the patent protection of drugs in developed countries, which creates a "gene divide". Therefore, under this background, "good technology" that could benefit human life and health has become "unattainable" technology that only a few people can afford. It is due to the fact that although technology is advancing, it cannot correspond to social fairness. We advocate scientific and technological progress in pursuit of efficiency and emphasis on resource allocation is in pursuit of fairness. Additionally, the continuous progress and development of Biobanks may lead to the birth of a "U-shaped society".

5. SUGGESTIONS

5.1. Take appropriate protective measures

It is vital to maintain the confidentiality of human genetic information. For the protection of human genetic information, two technical measures, de-identification or anonymization, can be mainly adopted, but both have their advantages and disadvantages.

First of all, de-identification is to remove identification marks by erasing sensitive personal information such as name, gender, and age. However, there is a risk of re-identification, which can be combined with other information to identify a particular individual. Therefore, researchers must take effective technical measures to prevent illegal access to genetic information, and carefully screen all possible foreseeable internal and external threats to human genetic information.

Secondly, although in most biomedical research, anonymization technology is a very important information security measure at present, which can provide powerful and effective means of identity information protection for the providers of human genetic information. However, it should be noted that although anonymization can cut off the connection with the human genetic information providers to a certain extent, complete anonymization makes the donor lose the possibility of withdrawing their consent, and the final research results and some accidental discoveries cannot be reported to donors, weakening the right of individual autonomy.

In short, both of them have some deficiencies to some extent. Therefore, researchers and institutions are required to continuously evaluate privacy risks during the research process, improve the protection level of technical measures, and timely repair information security vulnerabilities.

5.2. Choose the mode of informed consent carefully

The principle of informed consent is the core principle to protect human genetic information from being used illegally. However, it is important to consider carefully which informed consent model is more appropriate.

Informed consent models can be divided into several categories, such as Blanket consent represented by the Icelandic country, Broad consent represented by the United Kingdom, Dynamic consent represented by the Finnish country, Open consent, Specific consent, and Tiered consent.

The two informed consent modes of Blanket consent and Broad consent, belong to pre-authorization of the second use of genetic information in the future, [4]

greatly restricting the right of the human genetic information providers to make their own choices. In some cases, it may even be seriously contrary to their true intentions, for example, to disclose and share the genetic information of individuals who do not intend to share.

In contrast, Tiered consent has great advantages. That is to say, during the process of consent, participants will be listed with a list of different research types (e.g., heart disease, cancer) and they can elect for which research they consent to having their specimens and data used, and limit them to specific purposes. [5] It can be seen that this method can not only protect the autonomy of human genetic information providers, but also facilitate research institutions and researchers to use and analyze genetic information. In the end, it will help promote continuous progress in the field of biomedical research.

5.3. Give full play to the role of state organs

The protection of human genetic information by Chinese laws is still lacking, and the main approach is to regulate the relevant technical fields of human genetic information. Moreover, there is no special legislation for the protection of genetic information, and the relevant laws are not enough and systemic. Therefore, the relevant legislative organs must formulate specific and operable legal norms and policies to restrict the relevant research activities of Biobank, define the rights and obligations of scientific researchers and subjects, and improve the relevant procedures for the collection, use, and destruction of genetic information. In addition, the government guides and supervises the subjects to sign relevant agreements to protect the samples and genetic information, so as to protect the subjects' legal rights and interests.

5.4. Conduct rigorous ethical review

Give full play to the supervisory role of ethical review committees at all levels to conduct comprehensive reviews on the initiation and development of biomedical research activities. It includes rigorous review of the appropriateness of the protective measures adopted for human genetic information, the ethical rationality involved in biomedical research projects, and the substantive content of relevant documents such as informed consent documents. Ethics reviewers must undergo rigorous and systematic education and training, conduct in-depth study of relevant laws and regulations on human genetic information, fully respect the rights of human genetic information providers, continuously improve the scientific nature of ethics review, and always maintain absolute independence.

6. CONCLUSION

Nowadays, it has transitioned from the development stage of the traditional Biobank to the modern stage, and

has changed from focusing on "Biobank" to focusing on "biological data". As an important guarantee for advancement of medical technology and prevention and control of major diseases, the Biobank should surely be further deepened and expanded. But at this time, the issue of protection of human genetic information will become more prominent.

Therefore, it is necessary for us to make joint efforts, adopt a feedforward control method, and take different countermeasures for different risks that human genetic information may face, so as to resolve the dilemma of human genetic information protection. The ultimate goal is to achieve the balance between the smooth development of biomedical research and the protection of civil rights.

REFERENCES

- [1] W. G. Jürgen, S. Jürgen, "Do we need a biobank law?", *Journal of Laboratory Medicine*, vol. 43, no. 6, 2019, pp. 311-315. DOI: <https://doi.org/10.1515/labmed-2019-0065>
- [2] J. C. Stevenson, The century of the gene, *American Journal of Human Biology*, vol. 14, 2002, pp. 287-287. DOI: <https://doi.org/10.1002/ajhb.1144>
- [3] C. T. Fan, T. H. Hung, C. K. Yeh, "Taiwan Regulation of Biobanks", *Journal of Law, Medicine & Ethics*, vol. 43, no. 4, 2015, pp. 816-826. DOI: <https://doi.org/10.1111/jlme.12322>
- [4] M. J. Taylor, T. David, "Issues in Protecting Privacy in Medical Research Using Genetic Information and Biobanking: The PRIVILEGED Project", *Medical Law International*, vol. 10, no. 4, 2010, pp. 253-268. DOI: <https://doi.org/10.1177/096853321001000401>
- [5] M. A. Rothstein, B. M. Knoppers, H. L. Harrell, "Comparative approaches to biobanks and privacy", *The Journal of Law, Medicine & Ethics*, vol. 44, no. 1, 2016, pp. 161-172. DOI: <https://doi.org/10.1177/1073110516644207>