United States Government Accountability Office

Report to Congressional Committees

April 2025

HUMAN GENOMIC DATA

HHS Could Better Track Use of Foreign Testing Entities and Strengthen Oversight of Security Measures

GAO-25-107377

GAO

Highlights

Highlights of GAO-25-107377, a report to congressional committees.

April 2025

HUMAN GENOMIC DATA

HHS Could Better Track Use of Foreign Testing Entities and Strengthen Oversight of Security Measures

Why GAO Did This Study

Research and data on the human genome enable better ways to diagnose and treat diseases such as cancer. However, ODNI and others have warned of national security and other risks to Americans' genomic data. In February 2024, the President signed Executive Order 14117 to prevent access to Americans' bulk sensitive personal data, including personal health and human genomic data, and U.S. government-related data, by countries of concern. These countries, identified by the Department of Justice as directed by the Executive Order, are China, Russia, Iran, North Korea, Cuba, and Venezuela.

Congress included a provision in statute for GAO to review the risks and security measures to protect U.S. human genomic information. This report assesses (1) the risks and HHS's efforts to mitigate them, (2) HHS's and selected funding recipients' tracking of the use of genetic services from entities with ties to countries of concern, and (3) the data management and security policies and procedures that protect large-scale human genomic repositories.

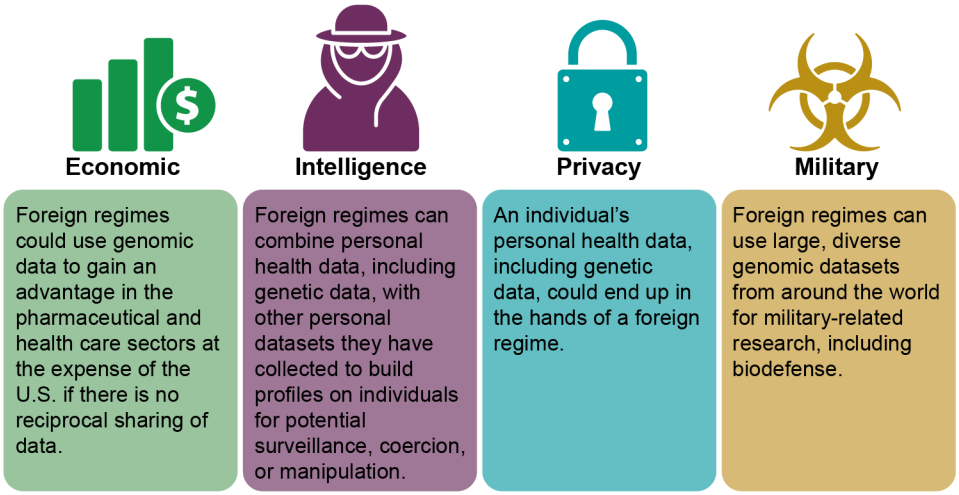
GAO reviewed documents and interviewed officials from ODNI and from HHS operating divisions involved in national security and genomics research or testing. GAO interviewed five subject matter experts who were selected to provide a mix of perspectives, based on their published works. GAO also interviewed eight HHS funding recipients selected based on the amount of funding from HHS for human genomic-related research. The selected recipients include a mix of research institutions, universities, and hospitals.

What GAO Found

A genome is the complete set of an organism's genes—all the information needed to build and maintain an organism (human or nonhuman) throughout its life. Genetic testing of a person's genome has multiple uses, such as diagnosing disease and identifying gene changes that may increase the risk of disease or that could be passed on to children. Within the Department of Health and Human Services (HHS), the National Institutes of Health (NIH) funds genomic research, the Centers for Disease Control and Prevention (CDC) conducts genomic research, and the Centers for Medicare & Medicaid Services (CMS) pays for medically appropriate genetic testing. In addition, NIH and CDC have repositories for researchers to collect and store genomic data for future research.

Foreign regimes in certain countries of concern pose risks to Americans' genomic data, according to the Office of the Director of National Intelligence (ODNI), other federal agencies, and selected experts, but HHS has not fully implemented mitigation measures. In 2021 and 2022, ODNI issued public warnings on the economic, intelligence, privacy, and military risks of Americans' genomic information being collected by foreign governments, noting China as having the motivation and capability to collect such information (see figure).

Risks to Americans' Human Genomic Information



Source: GAO analysis of Office of Director of National Intelligence and Department of Justice documents. (data); Krupal/valterz/stock.adobe.com (images). | GAO-25-107377

HHS officials described strategies to mitigate risks to genomic data through existing efforts, including the agency's policy to safeguard the acquisition of mission-critical products, materials, and services that it uses or funds through certain awards, including research grants. However, the HHS Office of National Security (ONS) has not implemented all elements of this policy. Specifically, ONS is required to, among other things, develop and share risk assessment standards and training for operating divisions, such as NIH and CDC, to apply when reviewing grants and mission-critical acquisitions. However, limited resources and differing funding priorities among HHS operating divisions have delayed

What GAO Recommends

GAO is making four recommendations, one to HHS ONS, two to NIH and one to CDC. Specifically:

- HHS ONS should develop and disseminate training and guidance on supply chain risk assessment standards that enable operating divisions to implement effective risk management for genomic data security while maintaining a focus on their core missions.
- NIH should begin systematically tracking the extent to which intramural and extramural researchers use genetic services provided by entities with ties to countries of concern.
- NIH should develop and implement procedures to proactively and comprehensively monitor researcher compliance with data management and security measures for human genomic data.
- CDC should develop and implement procedures, across all its centers that maintain restricted-access repositories with human genomic data, to proactively and comprehensively monitor researcher compliance with data management and security measures.

View GAO-25-107377. For more information, contact Candice N. Wright at WrightC@gao.gov.

efforts. Without risk assessment standards and training, operating divisions and HHS leadership are less equipped to apply the policy to grants and acquisitions related to human genomic information.

HHS officials and five selected funding recipients GAO spoke with described mostly using domestic genetic testing entities for research and treatment of patients. NIH officials described ways the agency monitors researchers' use of foreign entities, such as reviewing and approving requests to add foreign components to grant awards. However, NIH does not systematically track the use of foreign entities, in part because it collects limited information about genetic services for research. For work conducted by NIH's internal researchers, agency officials attributed this to limitations in NIH's procurement system, such as its inability to distinguish between funding for genetic services and funding for other research and laboratory services. NIH also did not have a code in its database of awards to track whether awards to researchers at external entities involved genetic services. NIH officials described measures they could implement to overcome these limitations, such as collecting more granular data from funding recipients on purchases from foreign entities. While researchers may mostly use domestic entities for genetic services, having information on researchers' use of foreign entities may allow NIH to help inform HHS decisions on how to restrict access to Americans' bulk genomic and other sensitive personal data by countries of concern.

NIH and CDC maintain repositories of genomic data and require researchers generating or using these data to follow data management and security measures. For example, pursuant to NIH policy researchers should strip data of personal identifiers according to specified regulations. The agencies also restrict access to certain repositories and investigate data management or security violations. NIH provided GAO with information on various types of confirmed violations between July 2018 and May 2024 (see table). NIH identified these violations through its review of researcher progress updates, researcher self-reports, or whistleblowers.

Examples of NIH Genomic Data Management and Security Violations, July 2018 – May 2024
Research conducted outside of the approved request
Security breach, such as compromised servers
Research outside of secondary use limitations, such as for-profit research
Data accessed by unapproved users
Data identifiers were not removed

Source: GAO analysis of genomic data management violations from the National Institutes of Health (NIH). | GAO-25-107377

However, NIH has not developed or implemented procedures to comprehensively monitor researchers' compliance with data management and security requirements. For example, NIH expects researchers who submit data to its repositories to certify that they meet certain data management and security requirements and restrict physical access to servers storing genomic data. In addition, NIH does not proactively audit the implementation of specific data security requirements, such as data encryption. Similarly, CDC officials stated that not all its centers that have repositories conduct oversight of whether funding recipients comply with data management or security measures for safeguarding health data, including genomic data, as stated in its policies. Therefore, NIH and CDC may be missing violations related to the data management and security of human genomic data that go unreported by researchers. Such violations could leave Americans' genomic data at risk of improper use by foreign regimes in countries of concern.

Contents

Abbreviations

CDC	Centers for Disease Control and Prevention
CLIA	Clinical Laboratory Improvement Amendments
CMS	Centers for Medicare & Medicaid Services
DAC	Data Access Committee
DNA	deoxyribonucleic acid
DOJ	Department of Justice
FBI	Federal Bureau of Investigation
GDS	Genomic Data Sharing
GINA	Genetic Information Nondiscrimination Act of 2008
HHS	Department of Health and Human Services
HIPAA	Health Insurance Portability and Accessibility Act of 1996
NHANES	National Health and Nutrition Examination Survey
NIH	National Institutes of Health
ODNI	Office of the Director of National Intelligence
ONS	Office of National Security
PHI	protected health information
SEED	Study to Explore Early Development

U.S. GOVERNMENT ACCOUNTABILITY OFFICE
441 G St. N.W.
Washington, DC 20548

April 30, 2025

Congressional Committees

The Human Genome Project completed the effort to generate the first sequence of the human genome (the complete set of a person's genetic information) in April 2003. All of the data generated by the Human Genome Project were deposited into public databases and made freely available to scientists around the world, with no restrictions on its use or redistribution. According to the National Institutes of Health (NIH), an operating division within the Department of Health and Human Services (HHS) that led the project, genome-based research is enabling better ways to diagnose and treat diseases such as cancer and diabetes. This research has the potential to spark tremendous economic growth—for example, by supporting new, high-wage jobs and building large datasets that can be used to develop emerging technologies such as gene therapy. Further, a May 2021 report commissioned by the American Society of Human Genetics estimated that the human genetics and genomics sector had a total economic impact of $265 billion on the U.S. economy.[1]

HHS, through NIH and its other operating divisions, is a key source of funding for genome-based research and the use of genetic testing to treat patients. For example, NIH announced in March 2022 that the NIH-supported *All of Us* Research Program had released its first dataset of nearly 100,000 whole genome sequences. The program aims to enroll a diverse group of at least 1 million persons in the United States to accelerate biomedical research. NIH policy calls for the broad and responsible sharing of genomic research data to facilitate the translation of research results into knowledge, products, and procedures that improve human health.

However, the Office of the Director of National Intelligence (ODNI) and the Federal Bureau of Investigation (FBI) have warned of national security and other risks to Americans' genomic data, such as use of the data by foreign regimes to build profiles of individuals for potential surveillance, coercion, or manipulation. ODNI's annual threat assessments for 2023 and 2024 noted that China has collected U.S. health and genomic data

[1]The American Society of Human Genetics, *The Economic Impact and Functional Applications of Human Genetics and Genomics* (Rockville, MD: Teconomy Partners LLC, 2021), 10.

through its acquisitions and investments in U.S. companies and the 2024 assessment notes that China's large volume of such data potentially positions it to lead in precision medicine, which involves using information about a person's genes and other factors to determine specific treatments.[2] In February 2024, in response to risks to national security, the President issued Executive Order 14117 on preventing access by countries of concern to Americans' bulk sensitive personal data, including genomic data and U.S. government-related data.[3]

The Consolidated Appropriations Act, 2023 includes a provision for GAO to review the risk of countries of concern obtaining U.S. human genomic information and the security measures to protect those data.[4] This report assesses (1) the risks associated with countries of concern obtaining Americans' genomic information that intelligence agencies and selected subject matter experts have identified and HHS efforts to mitigate those risks, (2) the extent to which HHS and its funding recipients track the use of genetic services from entities with ties to countries of concern and factors that may affect the use of foreign versus domestic entities, and (3) the data management and security policies and procedures that HHS and its funding recipients have established to protect large-scale human genomic repositories from being used against U.S. national security interests.

To address these objectives, we reviewed documents and interviewed or obtained written responses to questions from ODNI, FBI, and selected HHS offices and operating divisions. These HHS entities include the Office of National Security (ONS), NIH, the Centers for Disease Control and Prevention (CDC), and the Centers for Medicare & Medicaid Services

[2]Office of the Director of National Intelligence, *Annual Threat Assessment of the U.S. Intelligence Community* (Washington, D.C.: Feb. 6, 2023), 27; and *Annual Threat Assessment of the U.S. Intelligence Community* (Washington, D.C.: Feb. 5, 2024), 9.

[3]Exec. Order No. 14117, *Preventing Access to Americans' Bulk Sensitive Personal Data and United States Government-Related Data by Countries of Concern*, 89 Fed. Reg. 15421 (Mar. 1, 2024). The Executive Order defines a country of concern as "any foreign government that, as determined by the Attorney General pursuant to section 2(c)(iii) or 2(f) of this order, has engaged in a long-term pattern or serious instances of conduct significantly adverse to the national security of the U.S. or the security and safety of U.S. persons, and poses a significant risk of exploiting bulk sensitive personal data or U.S. government-related data to the detriment of the national security of the United States or the security and safety of United States persons, as specified in regulations issued by the Attorney General pursuant to section 2 of this order."

[4]Pub. L. No. 117-328, div. FF, tit. II, subt. C, ch. 3, § 2325, 236 Stat. 4459, 5768-69. (Dec. 2022).

(CMS). We selected HHS entities that have a role in addressing national security risks or that are primary sources of funding for or collaboration on genomics research and the use of genetic testing to treat patients. In addition, we interviewed nongeneralizable samples of five non-government subject matter experts in the national, health, and data security fields and representatives of eight HHS funding recipients, including universities, research institutes, and a hospital, about their use of genetic services. For this report, we defined genetic services to include a genetic test, genetic counseling (including obtaining, interpreting, or assessing genetic information), or genetic education.[5] See appendix I for more information on our objectives, scope, and methodology.

We conducted this performance audit from February 2024 to April 2025 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Background

Human Genomic Information

A genome is the complete set of an organism's genes—all the information needed to build and maintain an organism (human or nonhuman) throughout its life. Genetic testing of a person's genome has multiple uses, such as diagnosing disease and identifying gene changes that may increase the risk of disease or that could be passed on to children. In addition, researchers use genetic testing to expand their understanding of the human body, health, and disease. According to NIH, genomic medicine is an emerging discipline that involves using genomic information about an individual as part of their clinical care and has already made an impact in cancer screening and treatment, infectious disease, and other fields of medicine.

Within HHS, NIH funds genome-based research, CDC conducts genome-based research and public health surveillance, and CMS pays for

[5]Pub. L. No. 110-233, § 201(6), 122 Stat. 881, 907 (2008) (codified at 42 USC § 2000ff(6)). Section 2325 of the Consolidated Appropriations Act, 2023 uses the definition of genetic services found in section 201(6) of the Genetic Information Nondisclosure Act of 2008 (GINA). Pub. L. No. . Pub. L. No. 117-328, § 2325(a), 236 Stat. at 5768.

medically appropriate genetic testing to treat beneficiaries of its programs. Specifically:

- NIH funds extramural genomic research at external organizations, such as universities and hospitals, and intramural research in its own laboratories and clinics.
- CDC conducts intramural research, public health surveillance, and public health response activities that involve collecting, sequencing, analyzing and storing human genomic information.
- NIH and CDC have established repositories of human genomic information from past studies that researchers can use for ongoing projects. For example, GenBank ®, one of NIH's oldest data repositories, began in 1982 and contains DNA sequences for more than 581,000 named organisms as of November 2024.
- Under the Medicare program, CMS pays medically necessary claims for clinical diagnostic laboratory tests, including genetic tests, that have been ordered and used promptly by a health care provider who is treating a program beneficiary. State Medicaid programs may also pay for medically necessary laboratory tests.
- CMS administers the Clinical Laboratory Improvement Amendments (CLIA) program for ensuring the quality of laboratory testing (except some research tests) performed on human specimens.[6] Domestic and international facilities must obtain certification from the program if they (1) meet the definition of a "laboratory" under CLIA and (2) conduct testing on materials from human specimens collected in the United States, regardless of whether a test service is billed to Medicare.[7] According to CMS, genetic tests are considered high complexity tests and are therefore subject to stringent requirements under CLIA.

The Office of National Security (ONS), within the Office of the Secretary at HHS, is responsible for integrating intelligence and security information into HHS policy and operational decisions; assessing, anticipating, and warning of potential security threats; and providing policy guidance on and managing implementation of the department's security, intelligence, and counterintelligence programs.

[6]See Pub. L. No. 100-578, § 2, 102 Stat. 2903 (codified as amended at 42 U.S.C. § 263a).

[7]See 42 C.F.R. pt. 493.

Federal Policies and Requirements Related to Human Genomic Information

On February 28, 2024, the President signed Executive Order 14117 to prevent access to Americans' bulk sensitive personal data—including personal health data and human genomic data—by countries of concern when such access would pose an unacceptable risk to national security.[8] The Executive Order directs the Attorney General to, among other things, identify countries of concern. In January 2025, the Department of Justice issued a final rule implementing portions of the Executive Order, including identifying countries of concern as China, Russia, Iran, North Korea, Cuba, and Venezuela.[9]

The Executive Order also includes a requirement for the Attorney General to issue regulations that prohibit or otherwise restrict American individuals or entities from engaging in certain transactions involving bulk sensitive personal data or U.S. government-related data that has been determined by the Attorney General to pose an unacceptable risk to national security. The Executive Order and Department of Justice final rule except official business of the U.S. government from the prohibition of or restriction on certain transactions. The Department of Justice final rule further excepts transactions conducted pursuant to a grant, contract, or other agreement entered into with the U.S. government from restrictions on transactions covered by the final rule. In addition, the Executive Order directs the Secretary of HHS to, among other things, consider taking steps to prohibit the provision of assistance that enables access by countries of concern to Americans' bulk sensitive personal data, including human genomic data, or to impose mitigation measures with respect to such assistance.

Several HHS requirements also apply to human genomic information:

- Under the Health Insurance Portability and Accountability Act of 1996 (HIPAA), HHS has established rules to protect the privacy and security of certain individually identifiable health information, called protected health information (PHI), which could include human

[8]Exec. Order No. 14117, *Preventing Access to Americans' Bulk Sensitive Personal Data and United States Government-Related Data by Countries of Concern*, 89 Fed. Reg. 15421 (Mar. 1, 2024).

[9]Preventing Access to U.S. Sensitive Personal Data and Government-Related Data by Countries of Concern or Covered Persons, 90 Fed. Reg. 1636 (Jan. 8, 2025) (to be codified at 28 C.F.R. pt. 202).

genomic information.[10] According to the HHS Office for Civil Rights, which implements HIPAA, the HIPAA Privacy Rule only applies to covered entities, such as health plans and most health care providers and their business associates.[11] The HIPAA Privacy Rule permits covered entities to use and disclose PHI for research with the written authorization of an individual or without authorization under a waiver by an Institutional Review Board or Privacy Board.[12] An entity conducting research that is not a covered entity, such as NIH, is not subject to the HIPAA Rules. In addition, the HIPAA Privacy Rule governs the de-identification of data, including demographic data that could be used to identify an individual, such as name, address, birth date, and Social Security number. The HIPAA Privacy Rule specifies two ways to de-identify information: (1) expert determination that risk of identification is very small and (2) removal of specified identifiers of the individual and no actual knowledge that residual information can identify the individual.[13] Once PHI, such as genomic information, is de-identified in accordance with these requirements, it is no longer covered by this rule.[14]

- NIH's Genomic Data Sharing (GDS) Policy promotes sharing, for research purposes, of large-scale human and non-human genomic

[10]The HIPAA Rules refer to the Privacy, Security, Enforcement, and Breach Notification Rules—the regulations that implement HIPAA and set out requirements governing covered entities and business associates' disclosure, use, maintenance, and transmission of PHI. Pub. L. No. 104-191, Title II, Subtitle F, 110 Stat. 1936, 2021 (Aug. 21, 1996) as amended, 45 C.F.R. pts. 160 and 164 subpts. C (Security Rule), and E (Privacy Rule). PHI is individually identifiable health information and includes information that (1) is created or received by a covered entity; (2) relates to the past, present, or future physical or mental health condition of an individual; the provision of health care to an individual; or the past, present, or future payment for the provision of health care to an individual; and (3) identifies the individual or presents a reasonable basis to believe that the information can be used to identify the individual. PHI includes genetic information, and the Privacy Rule defines genetic information to include a genetic test, which is defined, generally, as an analysis of human DNA, RNA, or chromosomes that detects genotypes, mutations, or chromosomal changes. 45 C.F.R. § 160.103.

[11]45 C.F.R. § 164.502. HIPAA-covered entities include health plans, health care providers who conduct certain transactions electronically, and health care clearinghouses. Business associates are third parties that (1) create, receive, maintain, or transmit PHI on behalf of a covered entity for a covered function; or (2) provide certain services to or for a covered entity that involve the disclosure of PHI. 45 C.F.R. § 160.103.

[12]45 C.F.R. § 164.514(i).

[13]45 C.F.R. § 164.514.

[14]45 C.F.R. § 164.502(d)(2). Health information that has been de-identified in accordance with 45 C.F.R. §164.514 is no longer considered individually identifiable health information and the Privacy Rule requirements do not apply.

data generated from NIH-funded research.[15] The policy includes expectations for, among other things, de-identification of human genomic data submitted to NIH-designated data repositories and the informed consent of research participants for future research uses of their data.[16] The policy states that the informed consent under which the data or samples were collected determines whether the data should be available through an open or restricted access repository. Whereas an open access database is publicly available to anyone, a restricted access repository is available to researchers only after NIH's review of their proposed research use.

- CDC has established a policy to manage and ensure public access to federally funded public health data, including human genomic information, regardless of whether the data were collected as part of research.[17] The policy applies to the intramural and extramural collection of public health data. For example, it calls for a data management plan that describes the mechanisms for providing access to and sharing of data, including provisions for the protection of privacy, confidentiality, and security. The policy allows for access restrictions including when releasing data would risk disclosing proprietary or confidential information or compromising national security or law enforcement interests.

[15]National Institutes of Health, *NIH Genomic Data Sharing Policy*, NOT-OD-14-124 (Aug. 27, 2014). The policy applies to all NIH-funded research that generates large-scale human or non-human genomic data and the use of these data for subsequent research. Compliance with the policy is a special term and condition in applicable grant and contract awards. The *NIH Grants Policy Statement*, which generally serves as the terms and conditions for NIH grant awards, also includes policy requirements related to the use and sharing of genomic data and research.

[16]According to the Genomic Data Sharing Policy, researchers should de-identify human genomic data submitted to genomic databases according to the HHS Regulations for the Protection of Human Subjects to ensure that the identities of research subjects cannot be readily ascertained. In addition, researchers should strip the data of identifiers in accordance with the HIPAA Privacy Rule.

[17]Centers for Disease Control and Prevention, *Policy on Public Health Research and Nonresearch Data Management and Access,* CDC-GA-2005-14 (updated Jan. 26, 2016).

ODNI, FBI, and Experts Have Identified Risks to Genomic Data, but HHS Has Not Fully Implemented Mitigations

Countries of concern pose privacy, economic, and other national security risks to Americans' genomic data, according to ODNI, other federal agencies, and experts we interviewed. HHS officials described strategies to mitigate risks to human genomic data, but HHS has not implemented all elements of its policy to safeguard the acquisition of mission-critical products, materials, and services used by the department or funded through research grants. As a result, HHS operating divisions and leadership are less equipped to apply the policy to grants, cooperative agreements, and acquisitions related to human genomic information, such as genetic testing for NIH-supported research, where applicable.

ODNI and Others Have Identified Privacy, Economic, Intelligence, and Military Risks to Americans' Genomic Data

ODNI issued public warnings in February 2021 and January 2022 that highlighted the risks of Americans' genomic information being collected by foreign governments, particularly noting China as having the motivation and capability to collect such information. According to ODNI, this collection can occur through both legal and illegal means. For example, legal means can include investing in U.S. firms that handle sensitive health care and other types of personal data and forming research partnerships with U.S. hospitals, universities, and other organizations. In contrast, illegal means include theft of research, cybercrime, and theft of personal health information from U.S. companies.

In its February 2021 warning, ODNI cited examples of China's legal and illegal means of collecting genomic and other health information, such as the acquisition of U.S. genomics companies by Chinese firms in 2013 and 2015 and the involvement of two individuals based in China in the hack of a U.S.-based health insurer and three other U.S. companies in 2015. In addition, ODNI stated that China views bulk personal data, including health care and genomic data, as a strategic commodity to be collected and used for its economic and national security priorities. Collecting and analyzing large genomic datasets from diverse populations can help foster new medical discoveries and cures that can have substantial commercial value.

According to ODNI, foreign regimes can combine personal health data, including genetic data, with other personal datasets they have collected to build profiles on individuals. Moreover, according to Executive Order 14117 on preventing access by countries of concern to Americans' bulk sensitive personal data and U.S. government-related data, the combination of advances in technology and access to large datasets increasingly enable countries of concern to reidentify data, which may reveal Americans' health information that can be exploited.

ODNI has identified privacy, economic, intelligence, and military risks associated with foreign governments' collection of U.S. personal health data, including genomic data. See table 1 for a summary of these risks and examples cited by ODNI, other government sources, and subject matter experts we interviewed.

Table 1: Economic, Intelligence, Privacy, and Military Risks to Americans' Genomic Information

Risks cited by ODNI	Examples
Economic: Collection of large, diverse genomic datasets by foreign regimes and companies can boost their economic advantage in pharmaceutical and health care sectors at the expense of the United States when there is no reciprocal sharing of health data.	• According to the Office of the Director of National Intelligence (ODNI) and the Federal Bureau of Investigation (FBI), foreign countries can use U.S. human genomic data to develop new pharmaceuticals or precision medicine techniques, potentially displacing U.S. firms as global biotech leaders. • According to ODNI, the U.S. could become dependent on Chinese innovation and drug development to cure certain conditions, leading to a transfer of wealth and greater job opportunities in China. • According to the National Institutes of Health (NIH) and two experts, foreign countries may be able to leverage large, diverse genomic datasets from the U.S. to develop new pharmaceuticals and biotechnologies without providing the same opportunities for U.S. researchers.
Intelligence: Foreign regimes can combine personal health data, including genetic data, with other personal datasets they have collected to build profiles on individuals for potential surveillance, coercion, or manipulation.	• According to ODNI, foreign countries can use genomic information for surveillance or coercion of military, intelligence, or political officials and to act against dissidents located in the U.S. • According to ODNI and the Department of Justice (DOJ), foreign countries that use genomic information to better understand U.S. national security personnel vulnerabilities or to predict their behaviors could be more successful at recruiting such personnel for intelligence purposes. • One expert told us that foreign countries could identify undercover intelligence officers or agents through DNA left behind during sensitive operations, compromising U.S. covert activities.
Privacy: An individual's personal health data, including genetic data, could end up in the hands of a foreign regime.	• According to ODNI and FBI officials, foreign countries can use U.S. human genomic information to identify and take action against individuals in genetic subpopulations. For example, ODNI cited China's use of genetic analysis to further its repression of Uyghurs and other Muslim minority groups. In addition, a Chinese genomics company accessed U.S. persons health records and genetic data through partnerships with U.S. research and health care entities, according to ODNI. • One expert told us that public disclosure of genomic data could put individuals at risk of being stigmatized because of health conditions. • DOJ described the possibility that revealing that a political candidate had the genetic indicators for a debilitating health condition could impact the public's perception of that candidate or cause harm to their family.

Risks cited by ODNI	Examples
Military: Foreign regimes can use large, diverse genomic datasets from around the world for military-related research, including biodefense.	• According to ODNI and one expert, foreign countries can use genomic information to help advance dual-use technology with military and commercial application that could enable development of novel biological weapons. • Three experts stated that countries of concern could use U.S. human genomic information to harm ethnic or other subpopulations with genetically targeted pathogens. One of these experts described this threat as a futuristic scenario because of the costs and technical difficulties involved with implementing such a weapon. • NIH noted that there continues to be scientific debate about the potential for such scenarios.

Source: GAO analysis of ODNI and DOJ documents and statements from NIH and selected experts. | GAO-25-107377

Experts we interviewed described the potential for foreign countries or malicious actors to reidentify human genomic information. Two experts stated that foreign countries could use emerging technologies such as artificial intelligence to improve their reidentification capabilities. Additionally, in a 2018 study, researchers projected that 60 percent of searches on an individual's genomic information in large public repositories of human genomic information, such as GEDmatch and the 1000 Genomes Project, could result in a third cousin or closer match to the individual.[18] According to the study, repository users could theoretically identify an individual by combining search results with other demographic identifiers about the individual, such as age, sex, or ethnicity.

Experts also described commercial tactics China employs to obtain access to large, diverse human genomic datasets. Two experts described how Chinese businesses obscure their country of origin to obtain access to human genomic information—for example, by using front companies or frequent name changes. Moreover, according to ODNI and one expert we interviewed, China's national security laws allow for the government to compel private companies operating in the country to share data they have collected with the government. An FBI official told us that U.S. researchers may not conduct effective due diligence in evaluating their business partnerships, especially partnerships in China, creating opportunities for China to obtain U.S. human genomic information.

NIH officials told us that, despite potential risks to human genomic information, U.S. economic competitiveness is bolstered by sharing

[18]Yaniv Erlich, Tal Shor, Itsik Pe'er, and Shai Carmi, "Identity Inference of Genomic Data Using Long-Range Familial Searches," *Science* 362, no. 6415 (2018): 690-694. The researchers' projections were limited to individuals of European descent.

genomic data with both domestic and foreign researchers, as doing so helps facilitate scientific and medical discoveries that benefit the United States. According to NIH, restricting foreign researchers' access to U.S. genomic data places additional restrictions on domestic researchers, which can slow the pace of scientific and medical discoveries, creating opportunities for other countries to become world leaders in genomic research.

HHS Has Strategies to Mitigate Risks to Human Genomic Data but Has Not Fully Implemented Policies

Officials from NIH, CDC, CMS, and ONS described various strategies they have used to address risks to human genomic information, though some of the strategies are not specifically intended to address such risks. These strategies include the following:

- **NIH** stated that it consults with national security experts both for setting data sharing and access policies and practices for its restricted-access repositories of human genomic data, as well as for input on specific data access requests. In particular, NIH stated that it has consulted with the agency's Office of Research Services on specific data access requests and that this office can in turn consult with the broader national security, law enforcement, and intelligence communities to inform NIH decisions. NIH stated that it also monitors peer-reviewed scientific literature for new methods or capabilities that may pose privacy or other kinds of risks to human genomic information and other types of sensitive health information. Regarding NIH-funded human genome research, NIH officials stated that researchers must provide information to research participants on the risks of participating in a study, including potential privacy and reidentification concerns, and obtain oral or written verification from participants that they understand these risks.
- **CDC** officials stated they attend meetings with other HHS operating divisions to discuss concerns related to human genomic sequencing and participate in a National Security Council working group that develops strategies for protecting human genomic information. In addition, the officials stated that they participated in an interagency effort to provide risk information to the Department of Justice as part of the department's implementation of Executive Order 14117.
- **CMS** officials described strategies they apply as part of their implementation of the agency's CLIA laboratory certification

program.[19] According to CMS, under the program, test reports indicate the name and address of the laboratory where the test is performed, including if a test is performed at an international laboratory. In addition, CMS stated that it must be notified if the ownership or directorship of a laboratory changes after the initial application to the program. Separately, according to CMS, its Division of Strategic Information has begun to work directly with ONS to review information regarding foreign adversary access to genomic information and the type of technologies involved. Also, CMS officials stated that, as of January 2025, the Division of Strategic Information discusses national security risks associated with CLIA-certified labs located in countries of concern with ONS and participates in national security-related meetings and working groups organized by ONS.

- **ONS** officials said they frequently work with NIH and CDC, and also CMS, on issues related to genomic information, and seek to keep pace with fast-evolving genomic technologies. ONS officials explained that ONS relies on HHS operating divisions to approach ONS with requests to assess and recommend mitigations for national security risks. ONS officials told us that in fiscal years 2023 and 2024, ONS received 20 such requests from operating divisions regarding human genomic information. They told us that ONS recommended various mitigation strategies in response to these requests, such as operating divisions implementing stricter data access controls, enhanced vetting of foreign nationals, and terminating high-risk foreign partnerships. The officials cited various policies they draw on to vet foreign nationals before approving access to sensitive information, address insider threats to sensitive information, and conduct supply chain risk assessments. Additionally, ONS officials told us they conduct monthly meetings with operating divisions on supply chain risk assessments and that they conduct separate monthly meetings to brief operating division contracting officers. Officials from NIH and CDC confirmed that they attend such meetings.

In December 2022, HHS issued the Supply Chain Risk Management Policy to protect HHS's supply chain from threats to mission-critical products, materials, and services used by the department or its operating divisions. In December 2024, HHS issued a revised policy to, among other things, cover grants and cooperative agreements in addition to contracts. The policy establishes a process to identify, assess and

[19]Under the CLIA program, CMS regulates lab tests done on humans in the U.S., with some specified exceptions, and clinical laboratories are required to be certified by CMS before they can accept human samples for testing. See 42 U.S.C. § 263a, 42 C.F.R. pt. 493.

mitigate risks from foreign and other adversaries toward mission-critical products, materials, information, and services. According to ONS officials, examples of factors considered as part of supply chain risk assessments include company locations and foreign ownership, control, or influence.

However, ONS has not fully implemented its responsibilities under the department's policy. Under both the 2022 and 2024 versions of the policy, ONS is responsible for managing and overseeing the department's supply chain risk management program, and each HHS operating division must implement a program within the agency-wide program, unless otherwise delegated. The policy requires ONS to develop and share standards with operating divisions describing and incorporating supply chain risk assessments into mission-critical acquisitions and grants, in coordination with operating division program managers and to develop training programs, among other things. The standards and training could include guidance on how, if at all, HHS operating divisions can apply the policy to awards involving genomic information.

ONS officials stated that, as of January 2025, they had not fulfilled these responsibilities, in part because of limited resources, and that they expect to fully implement the policy by August 2025. They stated that insufficient staffing and funding hindered ONS's ability to finalize training materials and share risk assessment best practices with operating divisions. The officials told us they currently have about half as many staff as they would need to fully carry out their mission requirements, based on an internal workforce assessment conducted in 2024, and that they plan to address these challenges in part through process efficiencies.[20] They also stated that differing missions and funding priorities in HHS operating divisions, like NIH and CDC, led to delays and hindered implementation of effective risk management for genomic data security. For example, ONS officials noted that NIH is primarily focused on advancing health research rather than on security.

Because ONS has not developed or shared risk assessment standards or training, operating divisions and HHS leadership are less equipped to apply supply chain risk management to grants, cooperative agreements, and acquisitions related to human genomic information, such as genetic testing for NIH-supported research, where applicable. ONS officials stated that, despite the differing missions and funding priorities among operating divisions, their goal is to fully implement the supply chain risk

[20]We did not review the workforce assessment.

management program and consistently apply standards and procedures across the department. By developing standards and training that bridge differences in HHS operating divisions' priorities, ONS can enable more effective supply chain risk management for genomic information.

HHS Does Not Systematically Track Use of Foreign Genetic Services but Described Mostly Using Domestic Entities

NIH and CDC officials described mostly using domestic entities with no ties to countries of concern for genetic services. However, NIH officials stated that they do not systematically track the use of foreign entities for genetic services in extramural or intramural research. Selected extramural research funding recipients also stated that they mostly use domestic entities for research and in clinical settings to diagnose and treat patients, and they cited various factors they consider when determining where to procure genetic services, such as cost and speed of service.

NIH Conducts Limited Tracking of Foreign Genetic Services for Research

NIH funds both extramural and intramural research, which may involve conducting or procuring genetic services, such as genetic testing. NIH conducts some reviews when genetic services are conducted or procured from entities in foreign countries but does not systematically track use of these services (see table 2). For example, it does not monitor or report on the total amount of funding it provides for these services, whether to domestic or foreign entities.

Table 2: NIH Tracking of Foreign Genetic Services for Extramural and Intramural Research

Tracking for extramural research	Tracking for intramural research
• NIH reviews award applications and annual progress reports to ensure that applicants disclose foreign components. • NIH reviews and approves funding recipient requests to add foreign components to grants. • However, NIH officials said they were unable to describe the extent to which extramural researchers procure genetic services from countries of concern, because it does not have a code in its database of awards to systematically track this information.	• NIH's purchasing system stores information about genetic services under a generic code that includes other types of lab tests and services. • As a result, NIH officials could not tell us the extent to which intramural researchers and acquisition professionals procure genetic services, whether from domestic or foreign entities.

Source: GAO analysis of National Institutes of Health (NIH) information. | GAO-25-107377

Note: Consistent with the Genetic Information Nondiscrimination Act (GINA) of 2008, we defined genetic services to include a genetic test, genetic counseling (including obtaining, interpreting, or assessing genetic information), or genetic education. The NIH Grants Policy Statement defines a foreign component as "the performance of any significant scientific element or segment of the project outside of the U.S., either by the recipient or by a researcher employed by a foreign organization, whether or not grant funds are expended." NIH Grants Policy Statement (Apr. 2024).

NIH officials described ways the agency monitors extramural researchers' use of genetic services in foreign countries. For example, when NIH officials review narratives in extramural award applications and annual progress reports, they can check whether applicants fully disclosed all foreign components. NIH would generally need to request State Department clearance, prior to making the award, if it involves a foreign component—for example under a subcontract or subaward.[21] As outlined in the NIH Grants Policy Statement, NIH officials are required to review and approve funding recipient requests to add foreign components to their awards.[22]

However, NIH officials said they were unable to describe the extent to which extramural researchers procure genetic services from countries of concern or domestic entities with ties to countries of concern. NIH provided us with a list of 239 extramural awards with awardees or performance sites based in countries of concern that were active as of August 2024. However, NIH was not able to determine whether these awards included funding specifically for genetic services because it did not have a code in its database of awards to systematically track this information. NIH officials said, to track services provided by foreign entities, they would need to require that funding recipients separately report procurements from foreign entities to NIH. In our analysis of abstracts and NIH-designated spending categories for these awards, we identified 45 of these awards as potentially involving genetic services. Further, NIH could not provide information on subawards from awardees to foreign entities for genetic services. NIH officials stated that the agency does not have its own system to track subawards and instead relies on government-wide systems that awardees use to report information about

[21]For more information on State Department reviews of NIH awards with foreign components, see GAO, *Federal Research: NIH Could Take Additional Actions to Manage Risks Involving Foreign Subrecipients*, GAO-23-106119 (Washington, D.C.: June 14, 2023), 17-18.

[22]The NIH Grants Policy Statement defines a foreign component as "the performance of any significant scientific element or segment of the project outside of the U.S., either by the recipient or by a researcher employed by a foreign organization, whether or not grant funds are expended." Activities that may be significant include the use of facilities or instrumentation at a foreign site. NIH Grants Policy Statement (Apr. 2024).

subawards.[23] We previously reported that information on federal funds provided through subawards is not fully known because of limitations in the data provided in response to federal reporting requirements for subawards.[24]

For intramural research, NIH officials stated that their purchasing system does not have the ability to track the extent to which agency researchers request to procure genetic services from domestic or foreign entities. The officials stated that prior to purchasing genetic services, intramural researchers use an online marketplace, maintained by a third party, to conduct premarket research on genetic services offered by internal NIH facilities and by domestic and foreign entities. The researchers then work with acquisition professionals to ensure that the procurement adheres to all federal and agency regulations and policies.

According to NIH officials, the online marketplace contained 42 vendors from China who offered genetic services as of August 2024, but the officials did not know whether or how often agency researchers or acquisition professionals had used these vendors to procure genetic services because the platform lacks the functionality to collect this information. NIH's purchasing system stores information about genetic services under a generic code that includes other types of lab tests and services. As a result, NIH officials could not tell us the extent to which intramural researchers and acquisition professionals procure genetic services, whether from domestic or foreign entities.

[23]According to the Federal Funding Accountability and Transparency Act (FFATA) and implementing regulations and guidance, agencies are required to disclose certain information about federal awards that equal or exceed the micro-purchase threshold on a single public-facing, searchable website. In addition, award recipients are required to report specified information on first-tier subawards—with some exceptions—associated with these awards in the FFATA Subaward Reporting System (FSRS). The goal of the reporting is to increase transparency and publicly available information on federal spending. Since 2010, agencies have required award recipients to report such subaward information in FSRS, which is not public. USAspending.gov, the public-facing, searchable source of spending data includes data submitted by federal agencies and award recipients pursuant to FFATA, as amended, including data from government-wide reporting systems, such as the Federal Procurement Data System, the System for Award Management, and FSRS. Pub L. No. 109-282, 120 Stat. 1186 as amended by The Digital Accountability and Transparency Act of 2014, Pub. L. No. 113-101, 128 Stat. 1146 (codified as amended at 31 U.S.C. § 6101 note); 2 C.F.R. pt. 170; Federal Acquisition Regulation (FAR) parts 4.6 and 4.14.

[24]See GAO-23-106119, 3-5. Also see GAO, *Federal Research: Information on Funding for U.S.- China Research Collaboration and Other International Activities*, GAO-22-105313 (Washington, D.C.: Sept. 29, 2022), 7-8, 12-13.

CDC sponsors and maintains a database of biospecimens collected through the National Health and Nutrition Examination Survey (NHANES), which extramural researchers may apply to access, sequence, and use for research.[25] CDC officials stated that extramural researchers who are part of the Study to Explore Early Development (SEED) may also request to use or sequence DNA samples as part of the study, and the officials stated that there have been no requests to have samples sequenced by foreign entities as of January 2025.[26] In addition, CDC officials stated that the agency requires researchers to document the use of genetic services in agency agreements.

Officials from NIH, CDC, and selected extramural research funding recipients described mostly using domestic entities with no ties to countries of concern for genetic services. Both NIH and CDC officials stated that their intramural researchers generally use internal genetic services, such as the NIH Intramural Sequencing Center or researchers' own labs or external domestic sources, such as U.S.-based companies. For example, CDC's National Birth Defects Prevention Study (NBDPS) involves sequencing human genomic information. According to program officials, genomic sequencing services for NBDPS were completed by external domestic facilities, including at NIH. Nevertheless, NIH and CDC officials did not cite prohibitions on using foreign entities for genetic services. For example, NIH officials stated that they permit researchers to use the entities that are best equipped to meet the needs of the research project and therefore have no requirements for intramural researchers to consider using internal or domestic services before outsourcing to foreign entities.

Similarly, selected recipients of NIH extramural research funding stated that they typically do not use genetic services in or with ties to countries of concern. The eight funding recipients we interviewed stated that they have internal capabilities, and most stated that they use domestic entities if procuring services externally. Four of the eight funding recipients stated that they procure genetic services from foreign countries. Representatives from one funding recipient stated that some research projects conducted with HHS funding involve procuring services from China, due to their genomic sequencing needs. They stated that they have strict protocols in

[25]Through NHANES, the CDC's National Center for Health Statistics (NCHS) collects data about the health of adults and children in the United States, including health exams, laboratory tests, and dietary interviews for participants of all age groups.

[26]SEED began in 2007 and aims to learn more about autism spectrum disorder.

place for engaging all new vendors, which involve screening them to ensure they are not on federal lists of restricted entities and evaluating the risk level of a foreign partnership based on the sensitivity of the research data involved. The representatives said that they inform researchers of identified risks and how to mitigate them when procuring services from countries of concern.

Federal standards for internal control call for management to use quality information to achieve an entity's objectives, such as addressing inappropriate interference by foreign governments over federally funded research. While researchers may typically use domestic entities for genetic services, having additional, more detailed information on the extent to which intramural and extramural researchers use foreign entities may allow NIH to help further national security interests related to protecting human genomic information. For example, such information could help inform HHS implementation of Executive Order 14117, which seeks to restrict access to Americans' bulk genomic and other sensitive personal data by countries of concern. By using its internal systems to collect and report additional data on genetic services that it funds though intramural procurements and foreign components of extramural awards, NIH may be able to more easily report on the number and costs of projects involving the use of genetic services from countries of concern.

NIH faces limitations in collecting such information, such as its reliance on government-wide systems to track subawards. Nevertheless, NIH officials described measures they could implement to overcome these limitations, such as collecting more granular data from funding recipients on purchases from foreign entities.

HHS and Selected Funding Recipients Use Domestic Entities for Clinical Genetic Testing

According to CMS, it has certified labs in China under its CLIA program that conduct clinical genetic testing. However, CMS officials stated that federal law prohibits payments to providers in foreign countries when patients are insured by Medicare.[27] In part due to this prohibition, HHS officials and selected funding recipients described using domestic genetic services, located in the U.S., in clinical settings to diagnose and treat patients. CMS also noted that it can make payments to U.S.-based labs

[27]42 U.S.C. §1395y(a)(4) prohibits payment incurred for items or services provided outside of the United States except as provided in accordance with very limited exceptions. State Medicaid programs may also pay for medically necessary laboratory tests. According to CMS, there is no requirement for Medicaid programs to pay for any services outside of the country and payment for Medicaid services provided outside of the country may only be provided pursuant to Section 6505 of the Affordable Care Act.

that are owned by foreign entities as long as the payments are made to financial institutions with an address in the United States. According to CMS, all 5 percent or greater owners of a provider, including foreign owners, are required to be reported on the provider's Medicare enrollment application. CMS officials stated that CMS vets foreign owners in several ways, such as screening them for sanctions and criminal convictions.

NIH officials also stated that they typically procure clinical genetic testing services from domestic sources. In particular, according to NIH, its National Human Genome Research Institute typically uses CLIA-certified labs in the U.S. when procuring and analyzing the results of genetic testing for participants in its research projects, such as clinical trials. Moreover, CDC administers the Newborn Screening Quality Assurance Program, which helps assure the accuracy, reliability, and effectiveness of newborn screening tests, including tests to screen for genetic conditions. CDC officials stated that while Chinese and other foreign laboratories may participate in the program, doing so does not authorize foreign laboratories to screen newborns within the United States.

Five of the selected HHS funding recipients that use genetic services in clinical settings told us they rarely if ever use foreign entities to obtain these services because they have their own internal, CLIA-certified laboratories or outsource to CLIA-certified labs in the United States.

Selected HHS Funding Recipients Consider Cost, Speed, and Other Factors When Procuring Genetic Services

The eight funding recipients we spoke with stated that they consider various factors when deciding whether to procure genetic sequencing services from foreign or domestic sources. Seven funding recipients stated they consider the cost of genetic services, and four of these stated that domestic options tend to be more expensive. According to one funding recipient, researchers may prefer to use a foreign entity that provides a suite of services throughout the life cycle of a project because it is more cost-effective and convenient than dividing up the needs of a project among different foreign and domestic entities. Moreover, two funding recipients stated that they were only able to find certain genetic services needed for clinical research from foreign entities.

In contrast, four of the eight funding recipients stated that domestic entities may be able to perform tests and return results more quickly than foreign entities. Further, recipients stated that domestic entities are often faster to work with because researchers do not need to ship samples overseas or take time to develop contracts that meet foreign regulations.

Five funding recipients also identified legal requirements as a challenge of working with foreign entities. For instance, one funding recipient stated that working with foreign entities requires consideration of international laws and regulations, such as the European Union's General Data Protection Regulation, which do not need to be considered when working with domestic entities. Another funding recipient noted that, when they conducted research in India, they were required to use testing services within that country and comply with its specific laws.

NIH and CDC Require Researchers to Meet Security Measures for Human Genomic Data Repositories but Have Not Ensured Full Compliance

As a data security measure, 21 of the 28 repositories of human genomic data that NIH maintains or supports are restricted-access repositories. In April 2024, it limited access to those restricted-access repositories for entities in countries of concern. In addition, NIH identifies violations of data management and security provisions involving restricted-access genomic data and imposes penalties in response. Similarly, CDC maintains four repositories of human genomic data and has established measures to restrict access. However, NIH has not developed comprehensive researcher compliance mechanisms for the management and security of these genomic data in repositories, and CDC has not done so for all its programs' repositories.[28]

NIH Restricts Access to Certain Human Genomic Data Repositories and Has Limited Access by Countries of Concern

According to NIH officials, NIH maintains or supports 28 data repositories with de-identified, individual-level human genomic data for research use.[29] For example, the National Institute on Alcohol Abuse and Alcoholism maintains a repository with genomic data and other health records from more than 500,000 individuals. Under NIH's GDS policy, researchers who submit data to these repositories should de-identify the data according to standards set forth in HHS regulations.[30] For example, researchers

[28]For this report, we did not evaluate information- or cyber-security measures of the data repositories themselves.

[29]According to NIH, it also supports six repositories with de-identified, individual-level human genomic information but is not directly involved in their operation.

[30]NIH expects researchers to de-identify human genomic data submitted to genomic databases according to the HHS Regulations for the Protection of Human Subjects (the Common Rule) to ensure that the identities of research subjects cannot be readily ascertained. NIH also expects researchers to remove identifiers from the data according to the HIPAA Privacy Rule. The GDS policy contains protective measures in addition to the HIPAA Privacy Rule. For example, researchers are encouraged to obtain a certificate of confidentiality to prevent compelled disclosure of any personally identifiable information they hold.

should strip data of identifiers and assign random, unique codes with a corresponding key held by the researchers' institution to allow for reidentification of research participants, if necessary. Other security measures that NIH can implement for its data repositories with human genomic data include use of cloud platforms, continuous data monitoring, and more stringent researcher requirements (see text box).

***All of Us* Data Repository and Data Security Measures**

One of the 28 NIH-supported repositories with individual-level human genomic data is the *All of Us* Research Program. The program was formed to collect data, biological genomic samples, health records, and surveys from more than a million people living in the United States. The *All of Us* repository differs in policy and practice from other NIH-funded repositories.

Cloud platform. Unlike other repositories, which allow data to be downloaded and stored externally, the *All of Us* Research Program requires that researchers analyze data within a cloud platform. According to NIH officials, the ability of researchers to conduct analysis without removing data from the platform provides added protection against data mismanagement.

Continuous data monitoring. According to NIH officials, the program conducts continuous oversight of data security. This approach differs from other repositories, which emphasize annual or case-by-case reviews of research. NIH stated that the program conducts random compliance audits and that the cloud platform automatically alerts the program when unidentified or unauthorized researchers access data from outside the U.S. or move a large amount of data.

Researcher requirements. The program requires researchers to follow data use and dissemination policies that are not among the requirements of the Genomic Data Sharing Policy, according to NIH officials. For example, researchers may not publish or distribute research findings corresponding to fewer than 20 study participants, whereas the Genomic Data Sharing Policy does not include this specific stipulation. In addition, the program conducts formal risk assessments of participant data to minimize privacy risks.

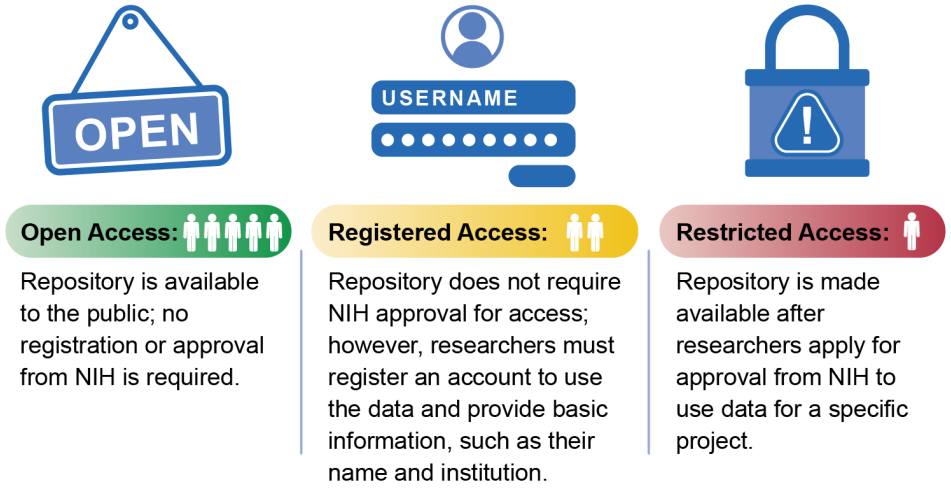
Source: GAO analysis of information from the National Institutes of Health (NIH). | GAO 25 107377

According to NIH, individual-level genomic data are restricted in 21 of the repositories it maintains or supports. NIH may restrict access to data for one or more of the following reasons:

- **Limitations on secondary research.** Data in the repository are restricted from subsequent use by agency policies or limitations in the informed consent provided by research participants. For example, if researchers did not obtain consent from research participants to share their data with for-profit entities, then these entities cannot use the data.
- **Sensitive data.** Certain data in the repository could reveal stigmatizing traits, illegal behaviors, or other information that could be used for discriminatory purposes or that may pose a high risk of re-identification.
- **Identifiable data.** The repository contains data that have been de-identified to regulatory standards but, according to NIH, have the potential to be reidentified using advanced scientific or technical means.

In addition to restricted-access repositories, NIH maintains or supports others that do not require the agency to prospectively control access (see fig. 1). For example, the National Center for Biotechnology Information allows researchers to make human and other organisms' DNA and RNA sequences publicly available through GenBank®, an open-access repository. According to NIH policy, genomic data can be shared in a repository without access controls if research participants have explicitly consented to share their de-identified data without restrictions.

Figure 1: Access Tiers of National Institutes of Health (NIH) Genomic Data Repositories



Source: GAO analysis of National Human Genome Research Institute information available at www.genome.gov/about-genomics/fact-sheets/Genomic-Data-Science (data and illustration). | GAO−25−107377

Under the GDS policy, researchers must apply to access data from any of NIH's 21 restricted-access repositories. Researchers must submit data access requests for all new research projects and reapply for access annually. According to NIH officials, the agency's Data Access Committees (DAC) consider whether a researcher's proposed use of restricted-access data is consistent with data use limitations, such as using the data for disease-specific research, and review for the potential to cause harm to individuals or groups through discriminatory practices.[31] The requirement to apply for access applies to both intramural and extramural research, regardless of whether researchers are direct

[31]NIH has 19 DACs that focus on different areas of biomedical research, such as Alzheimer's disease and cancer. Each DAC includes at least one chair and four or more NIH staff with science, policy, or bioinformatics expertise.

recipients of NIH funding, according to NIH officials. The officials also stated that DACs follow the same considerations to evaluate requests from domestic and foreign researchers.

According to NIH, researchers in countries of concern accounted for about 7 percent of the approximately 47,200 requests to use data from its restricted access repositories in fiscal year 2023. Researchers in China accounted for almost all these requests. DACs approved access requests from countries of concern at a similar rate as access requests from the United States and foreign countries other than those identified as countries of concern (around 78 percent).

In April 2024, NIH issued guidance to DACs directing them to reject new access requests from individuals and entities located in countries of concern. According to NIH officials, the agency issued this guidance after a small number of requests from entities in countries of concern raised doubts about the legitimacy of the request or the identities of the requestors. NIH officials provided several reasons for the policy to reject all new requests originating from countries of concern:

- **Administrative burden.** NIH stated that its reviews of data access requests from countries of concern are labor-intensive and that it is not administratively feasible to conduct such reviews given the volume of requests (an average of more than 2,500 per year in fiscal years 2018 through 2024).
- **Likelihood of rejections.** NIH stated that, even after its reviews, it would likely reject many data requests from countries of concern due to concerns about the requesting entity's ability to meet the terms and conditions of the GDS policy.
- **Anticipated effects of future limitations on data access.** NIH stated that ceasing approvals for new data access requests originating from countries of concern could safeguard against disruptions to research that policies in response to Executive Order 14117 may create. For example, according to NIH officials, NIH might decide to discontinue data access to researchers in a country of concern after initially approving access.

According to NIH officials, the agency has a temporary policy of continuing to review and approve, as appropriate, renewal requests for data access from researchers in countries of concern whose initial applications NIH had approved prior to its April 2024 guidance. NIH officials stated that the agency intends to update this policy in the near

future to align with national security directives, such as Executive Order 14117.

NIH Investigates Violations of Its Genomic Data Sharing Policy but Does Not Comprehensively Ensure Compliance

NIH's GDS policy establishes a framework for using and sharing large-scale human and nonhuman genomic data. This policy outlines expectations for researchers including for de-identifying data, developing data sharing plans, and meeting minimum data security requirements (see table 3).[32]

Table 3: NIH Genomic-Data-Security-Related Expectations for Researchers

Researchers subject to the policy	Security-related expectations and assurances for human genomic information
NIH-funded researchers who generate human genomic information	• Develop data sharing plans that specify to which NIH-hosted repository data will be submitted. • De-identify data to protect the identities of research subjects by stripping out identifiers as outlined in the Health Insurance Portability and Accountability Act's (HIPAA) Privacy Rule. • Obtain research participants' consent for data to be shared for secondary research use in an NIH repository. • Provide institutional certification of any limitations on the research use of the data, as expressed in research participants' informed consent.
Researchers who access human genomic information from an NIH-supported repository	• Use data only for research described in approved access requests. • Not attempt to re-identify individuals. • Agree to notify NIH of security breaches or unauthorized data sharing that may compromise data confidentiality within 24 hours of when an incident is identified. • Meet minimum data security requirements, including the use of strong authentication technology and physical security guidelines.

Source: GAO analysis of National Institute of Health's (NIH) Genomic Data Sharing Policy and related documents. | GAO-25-107377

Note: The Genomic Data Sharing (GDS) policy applies to both intramural and extramural researchers that receive NIH funding to generate human genomic data and sets terms of access for researchers who are not funded by NIH to data in NIH-maintained or supported repositories. Compliance with the policy is a special term and condition in an applicable NIH grant or contract award. The HIPAA Privacy Rule governs the de-identification of data, including demographic data that could be used to identify an individual.

[32]In January 2023, NIH implemented the NIH Policy for Data Management and Sharing, which applies to all research funded or conducted in whole, or in part, by NIH that results in the generation of scientific data, including genomic data. Under this policy, information expected in the genomic data sharing plans under the Genomic Data Sharing Policy is to be included in the Data Management and Sharing Plan rather than as a separate genomic data sharing plan. See, National Institutes of Health, *Implementation Changes for Genomic Data Sharing Plans Included with Applications Due on or after January 25, 2023*, NOT-OD-22-198 (August 2022).

According to NIH officials, researchers must adhere to the GDS policy if they generate or access genomic data from an NIH-supported repository.[33] If they fail to do so, NIH may take enforcement actions, such as developing a remediation plan or suspending access to genomic data.

Researchers may intentionally or inadvertently commit data management and security violations when submitting or accessing data in a restricted-access repository. DACs review researchers' required progress updates and will investigate if they have reason to believe a data management or security violation may have occurred, according to NIH officials. Once a potential data management or security incident is discovered or reported to a DAC, the committee begins a coordinated investigation.

NIH provided us with information on 40 confirmed violations between July 2018 and May 2024, of which 36 involved restricted data (see table 4). NIH identified violations through its DACs' review of annual renewal and close-out requests, researcher self-reports, or whistleblowers. According to NIH officials, none of these incidents involved countries of concern. The most common violations related to data management or security (with three or more occurrences) included researchers using data for purposes not stated in their data access requests, researchers sharing genomic data with unapproved users, and breaches of physical security, such as compromise of a server storing restricted-access data. In response to eight of these incidents, DACs imposed penalties, such as requiring remediation plans or suspending repository access. NIH was continuing to investigate nine incidents as of September 2024.[34] NIH requests that remediation plans be developed in partnership with the researcher and DACs and that they correct the violation and prevent it from occurring again, according to NIH officials. NIH often did not impose penalties on researchers who self-identified the violation and quickly corrected the issue.

[33]The GDS Policy applies to both intramural and extramural researchers that receive NIH funding to generate human genomic data and sets terms of access to human genomic data for researchers who are not funded by NIH to data in NIH maintained or supported repositories.

[34]NIH did not impose penalties in response to 21 of the data management or security violations. The additional two incidents were identified in response to an NIH error.

Table 4: Types of NIH Genomic Data Management and Security Violations, July 2018 – May 2024

Type of violation	Number of violations
No NIH database citation in final research publication[a]	10
Research conducted outside of the approved request	7
Security breach, such as compromised servers	7
Research outside of secondary use limitations	3
Data accessed by unapproved users	3
Data not de-identified	3
Other	7
Total	**40**

Source: GAO analysis of genomic data management violations from the National Institutes of Health (NIH). | GAO-25-107377

[a]We did not consider failure to cite an NIH database in a final research publication to be a data management or security violation but included it in our analysis to provide complete data on violations of NIH's Genomic Data Sharing Policy. Examples of "other" violations included NIH systems permitting data access to unapproved users and researchers' institutions storing data on servers that lacked proper security measures.

NIH officials said that to help identify these violations, DACs review the annual data access renewal requests and project close-out reports that approved genomic data users are required to submit as part of progress updates. NIH officials also said that DACs verify whether researchers have stayed within the scope of their research statements, cited NIH databases in publications, and included researchers without approval to use data.

Similarly, NIH program officers review annual performance progress reports submitted by NIH-funded researchers against a checklist with questions about implementation of researchers' plans to share data, including genomic data. According to NIH policy, program officers are expected to assess the adequacy of these plans.[35] Starting in 2025, researchers will be required to indicate how they have complied with their approved data management and sharing plan as part of the annual review, according to NIH officials. Further, in July 2024 NIH issued an update to security standards for restricted-access genomic data

[35]National Institutes of Health, NIH Policy for Data Management and Sharing, October 2020 (effective Jan. 25, 2023). The information expected in the genomic data sharing plan under the Genomic Data Sharing Policy is now provided as part of the broader Data Management and Sharing plan required by the NIH Policy for Data Management and Sharing. Compliance with the Genomic Data Sharing Policy is handled in accordance with the compliance and enforcement terms of the NIH Policy for Data Management and Sharing. NIH, NOT-OD-22-198.

repositories and their users and that this update went into effect in January 2025.[36] According to NIH officials, the update will provide greater assurance of the security of genomic data and opportunities for monitoring researcher compliance.

However, NIH has not developed comprehensive compliance mechanisms to ensure that researchers adhere to certain requirements of its GDS Policy and related policies. For example, NIH expects institutions that submit data to NIH's repositories to self-certify that they meet certain data management and security requirements and obtain research participants' informed consent. For researchers granted access to restricted genomic data, NIH does not proactively audit the implementation of minimum data security requirements, such as data encryption and restriction of physical access to servers, according to NIH officials. Rather, the officials said that NIH reviews institutions' implementation of these requirements on a case-by-case basis after it has reason to believe there has been a data management incident. NIH officials said comprehensive reviews of researcher adherence to data management and security policies would be challenging due to the large volume of genomic research the agency oversees. However, NIH already has systems in place, such as its reviews of data access renewal requests and researchers' progress updates, that it has not fully leveraged for oversight of compliance with the GDS and related policies.

Federal standards for internal control state that management should implement control activities through policies and procedures. Such control activities can help NIH provide reasonable assurance that it will achieve its objectives, such as responsible sharing of human genomic data, and respond to risks. Without procedures for proactively and comprehensively monitoring researcher compliance with data management and security requirements for human genomic data, taking into account agency resource limitations, NIH may be missing violations that go unreported by researchers.

[36]National Institutes of Health, Implementation Update for Data Management and Access Practices Under the Genomic Data Sharing Policy, NOT-OD-24-157 (July 2024).

CDC Restricts Access to Genomic Data but Has Not Ensured Compliance Across All Centers

CDC maintains four restricted-access repositories with de-identified human genomic information (see table 5).[37] Similar to NIH, CDC maintains a dual-access system for genomic information in its repositories, which includes open and restricted categories. According to CDC, it restricts access to data that can be used to identify people, CDC facilities, or other research institutions.

Table 5: CDC Repositories with Human Genomic Information

Repository Name	Associated Center
National Health and Nutrition Examination Survey (NHANES) Dataset	National Center for Health Statistics
Study to Explore Early Development (SEED) Dataset	National Center on Birth Defects and Developmental Disabilities
National Birth Defects Prevention Study (NBDPS) Dataset	National Center on Birth Defects and Developmental Disabilities
Birth Defects Study to Evaluate Pregnancy exposureS (BD-STEPS) Dataset	National Center on Birth Defects and Developmental Disabilities

Source: GAO analysis of information from Centers for Disease Control and Prevention (CDC). | GAO-25-107377

Under its Policy on Public Health Research and Nonresearch Data Management and Access, CDC expects researchers to develop data management plans prior to project initiation and make the resulting data available to CDC.[38] Researchers must comply with the requirements stated in the policy or may risk losing their access to genomic and other public health data. According to CDC officials, data should be transferred via secure channels and maintained data in locked containers or on encrypted, password-protected computer systems. CDC officials also stated that if a data breach or misuse of data occurs, researchers must report the incident to CDC for review.

Individual centers also maintain their own policies and practices for genomic data access and oversight of data security in some repositories:

- **National Health and Nutrition Examination Survey (NHANES).** According to CDC officials, to request access to NHANES restricted-use data, researchers must submit applications for review by

[37]According to CDC officials, the NHANES data repository uses an alternative to full de-identification.

[38]CDC's Policy on Public Health Research and Nonresearch Data Management and Access applies broadly to public health data, which can include genomic data. According to CDC officials, the policy ensures that data, which could include genomic data, are made available to intramural and extramural researchers subject to laws, ethical considerations, and other decision factors.

NHANES staff, the National Center for Health Statistics' (NCHS) Confidentiality Office, and the NCHS Research Data Center.[39] CDC officials also stated the application provides a framework to identify potential disclosure or reidentification risks and how the data will be used in secondary research. Once approved, researchers can only access restricted-use data at a secure location in the Research Data Center network. According to CDC officials, the Research Data Center has approved 12 requests to access NHANES restricted-use genomic data since fiscal year 2018. Data will not be released from a Research Data Center facility if the data do not match the approved research question, contain individual-level data, or could lead to reidentification of individuals, according to CDC guidance.

- **Study to Explore Early Development (SEED).** According to CDC officials, to obtain access to SEED genomic data, researchers must be part of the SEED Network and apply to a data sharing committee and meet security measures outlined in a cooperative agreement, including requirements for protecting de-identified genomic information.[40] According to CDC officials, there have been 11 requests to access SEED human genomic data since fiscal year 2018, but none from foreign researchers, including those from countries of concern. CDC officials said they monitor researcher compliance through site visits at least once every 2 years, which include review of researcher data security measures.
- **National Birth Defects Prevention Study (NBDPS).** Access to the two NBDPS repositories is restricted to researchers sponsored by an NBDPS principal investigator (lead researcher). CDC officials stated that, to date, NBDPS researchers have used genetic data from their local databases and have not requested data from the centralized repositories. Per the center's data sharing guidelines, the principal investigators are responsible for assuring researchers have signed the data use agreement on a yearly basis. In addition, CDC officials told us that any incidents of genomic data misuse or security breaches are reported for investigation and review. However,

[39]According to CDC officials, the Research Data Center is responsible for protecting confidentiality of survey respondents, study subjects, or institutions when providing access to restricted-use data for research purposes. For example, the center does not provide researchers with data that can directly identify participants, such as names, Social Security numbers, and addresses.

[40]According to CDC officials, the Data Sharing Committee is comprised of SEED network investigators and CDC program representatives. Any request to use DNA samples from the SEED biorepository must also be approved by the SEED Biomonitoring Committee, which includes laboratory and genetics subject matter experts from the network and CDC.

according to officials, the program and Prevention does not conduct researcher oversight on compliance with the provisions in the confidentiality and data use acknowledgement form stating that researchers should not share any genomic data with unapproved users or transfer data outside of secure channels.

As of September 2024, CDC had not identified any data security incidents related to genomic research at any of its centers. However, according to CDC officials, not all centers conduct oversight of whether researchers comply with data management or security measures for safeguarding health data, including genomic data. In particular, the National Center on Birth Defects and Developmental Disabilities relies on researcher self-certification of the data use agreement and expects researchers to self-report any incidents of data misuse or security.

CDC's policy on Public Health Research and Nonresearch Data Management and Access requires the monitoring of awardee compliance with the policy. In addition, federal standards for internal control state that management should implement control activities through policies and procedures. Such control activities can help CDC provide reasonable assurance that it will achieve its objectives—such as protecting health data, including human genomic data based on national security concerns—and respond to risks. Because the National Center on Birth Defects and Developmental Disabilities relies on researcher self-certification, CDC may be missing violations related to the data management or security of human genomic information that go unreported by researchers.

Conclusions

NIH, CDC, and CMS—three key HHS operating divisions involved in genomics research and genetic testing—have recognized and taken steps to address national security and other risks to Americans' genomic data. For example, NIH issued guidance to reject new requests for access to human genomic data repositories from individuals and entities located in countries of concern.

However, HHS agencies have not taken certain steps that would enable them to better address risks to Americans' genomic data. First, in part because of resource challenges that it plans to address through process efficiencies, HHS ONS has not fully implemented its responsibilities to develop and share supply chain risk assessment standards and develop training programs to implement the department's supply chain risk management process. As a result, HHS agencies and leadership are less

equipped to apply this process to research grants and acquisitions, including those related to human genomic information, if applicable.

In addition, NIH does not systematically track the use of genetic services including the extent to which it funds such services in foreign countries. The agency attributed this to limitations in its procurement system, such as its inability to distinguish between funding for genetic services versus other research and laboratory services. Having additional, more detailed information on the extent to which intramural and extramural researchers use foreign entities may allow NIH to help further national security interests related to protecting human genomic data.

Lastly, neither NIH nor CDC have fully developed or implemented procedures to comprehensively monitor researchers' compliance with data management and security requirements. For example, NIH expects researchers who submit data to its maintained or supported repositories to self-certify that they meet certain data management and security requirements. In addition, NIH does not proactively audit the implementation of specific data security requirements, such as data encryption. Similarly, CDC does not conduct oversight of whether researchers comply with its data management or security policy and procedures, which include measures for safeguarding health data, including human genomic data. Without proactively and comprehensively monitoring researcher compliance with data management and security requirements for human genomic information, NIH and CDC may be missing violations that go unreported by researchers. Such violations could leave Americans' genomic data at risk of improper use by foreign regimes in countries of concern.

Recommendations for Executive Action

We are making a total of four recommendations, one to HHS ONS, two to NIH, and one to CDC. Specifically:

The Secretary of HHS should direct that ONS develop and disseminate training and guidance on supply chain risk assessment standards that enable operating divisions to implement effective risk management for genomic data security while maintaining a focus on their core missions. (Recommendation 1)

The director of NIH should direct that NIH begin systematically tracking the extent to which intramural and extramural researchers use genetic services provided by entities with ties to countries of concern. (Recommendation 2)

The director of NIH should require the development and implementation of procedures to proactively and comprehensively monitor researcher compliance with data management and security measures for human genomic data. (Recommendation 3)

The director of CDC should direct CDC to develop and implement procedures, across all its centers that maintain restricted-access repositories with human genomic information, to proactively and comprehensively monitor researcher compliance with data management and security measures. (Recommendation 4)

Agency Comments

We provided a draft of this report to HHS, Department of Justice, and ODNI for review and comment. HHS provided written comments, which are reproduced in appendix II. HHS also provided technical comments, which we incorporated as appropriate. FBI and ODNI did not have comments on our draft.

In its comments, HHS concurred with our recommendation to ONS and described steps that ONS will take to address it. In addition, HHS stated that NIH and CDC concurred with our recommendations directed to those operating divisions. The written comments from HHS included a suggestion from NIH that our final report acknowledge its July 2024 update to security standards for restricted-access repositories. We agreed with the suggestion and added information about the updated standards to our report. NIH's update to security standards did not affect our finding or recommendation on NIH monitoring of researcher compliance with data management and security measures for human genomic data.

We are sending copies of this report to the appropriate congressional committees, the Secretary of Health and Human Services, the Attorney General, and the Director of National Security. In addition, the report is available at no charge on the GAO website at https://www.gao.gov.

If you or your staff have any questions about this report, please contact us at wrightc@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made key contributions to this report are listed in appendix III.

//SIGNED//

Candice N. Wright
Director, Science, Technology Assessment, and Analytics

List of Committees

The Honorable Bill Cassidy
Chair
The Honorable Bernard Sanders
Ranking Member
Committee on Health, Education, Labor and Pensions
United States Senate

The Honorable Tom Cotton
Chairman
The Honorable Mark Warner
Vice Chairman
Select Committee on Intelligence
United States Senate

The Honorable Brett Guthrie
Chairman
The Honorable Frank Pallone, Jr.
Ranking Member
Committee on Energy and Commerce
House of Representatives

The Honorable Rick Crawford
Chairman
The Honorable Jim Himes
Ranking Member
Permanent Select Committee on Intelligence
House of Representatives

Appendix I: Objectives, Scope, and Methodology

This report assesses (1) the risks associated with countries of concern obtaining Americans' genomic information that intelligence agencies and subject matter experts have identified and Department of Health and Human Services (HHS) efforts to mitigate those risks, (2) the extent to which HHS and its funding recipients track the use of genetic services from entities with ties to countries of concern and factors that may affect the use of foreign versus domestic entities, and (3) the data management or security policies and procedures that HHS and its funding recipients have established to protect large-scale human genomic databases from being used against U.S. national security interests.

The scope of our review included HHS agencies and operating divisions that provide funding for, oversight of, or have expertise in genetic services or national security risks to U.S. human genomic information: the National Institutes of Health (NIH), the Centers for Disease Control and Prevention (CDC), the Centers for Medicare & Medicaid Services (CMS), the Office of National Security (ONS), and the Office for Civil Rights (OCR). We selected these HHS operating divisions and offices because of their involvement in the funding and implementation of human genomic research and use of genetic testing or in the oversight of such activities. In addition, the scope of our objective on risks associated with countries of concern obtaining Americans' genomic information included the Office of the Director of National Intelligence (ODNI) and the Federal Bureau of Investigation (FBI). We selected ODNI and FBI based on their expertise on national security risks to U.S. human genomic information. As part of our overall methodology, we interviewed and obtained written responses to questions from all the agencies in our scope.

To assess risks associated with countries of concern obtaining Americans' genomic information and HHS' efforts to mitigate those risks, we reviewed documents such as public warnings from ODNI regarding national security risks to human genomic information and other public documents. These included ODNI's 2024 National Counterintelligence Strategy, Executive Order 14117 and the Department of Justice (DOJ) rule implementing portions of that executive order, as well as HHS's Supply Chain Risk Management Policy and training materials for implementing that policy.

We worked with a GAO research librarian to conduct a literature search of research on human genomic data, genetic services, and security. We conducted a broad search of materials published within the last 5 years, including scholarly articles and government reports. From these searches, we identified and selected relevant articles to include in our

review. We used the results of our literature review to inform our findings and selection of subject matter experts for interview.

We conducted interviews with officials from selected HHS operating divisions and offices, and we conducted interviews with five subject matter experts in the national, health, and data security fields to obtain additional context on the risks. The views of these experts could not be generalized. We identified subject matter experts through our literature search and background research and selected experts who could provide a mix of perspectives from across these fields based on their published works.

To assess the extent to which HHS and its funding recipients track the use of genetic services from entities with ties to countries of concern and factors that may affect the use of foreign versus domestic entities, we reviewed NIH's Grants Policy Statement regarding requirements for obtaining and executing grants. NIH also provided us with information on 239 awards related to human genomic sequencing with recipients or performance sites based in countries of concern. We identified 45 of those awards as relating to human genomic sequencing based on the spending categories and award abstracts. Similarly, NIH provided a list of Chinese entities in its Collaborative Research Exchange, and CMS provided information on foreign labs certified for CLIA testing. We analyzed this information to determine whether such entities could obtain funding from NIH research efforts. We also compared NIH's tracking of genetic services to federal standards for internal control. In particular, we determined that the principle that management should use quality information to achieve the entity's objectives was significant to this objective.[1]

We conducted semi-structured interviews with representatives from a sample of eight NIH funding recipients, including universities, research institutes, and hospitals, whose views could not be generalized, about their use of genetic services and any NIH requirements regarding their use of foreign entities for such services. We chose the eight funding recipients based on the relevance of their work to the human genomic field, as well as amount of funding NIH awarded them for work related to human genomic information, based on public funding information in NIH's Research Portfolio Online Reporting Tools Expenditures and Results

[1]GAO, *Standards for Internal Control in the Federal Government*, GAO-14-704G (Washington, D.C.: September 2014).

(RePORTER) repository. We analyzed this information to select a mix of funding recipients that had received high, medium, and low amounts of funding for genomic-related projects, according to NIH's spending categories within RePORTER. We conducted additional interviews with representatives from two industry groups, one representing researchers and the other representing hospitals, that had published information on genomic data security for their industries. Their views also could not be generalized.

To assess the data management or security policies and procedures that HHS and its funding recipients have established to protect large-scale human genomic databases, we reviewed the HIPAA Privacy Rule; HHS regulations for the protection of human subjects in research; NIH policies and guidance for sharing data, including the Genomic Data Sharing Policy and the guidance on working with entities in countries of concern; and CDC policies on data management and access. We also compared the data management or security policies and procedures at NIH and CDC to federal standards for internal control. In particular, we determined that the principle that management should implement control activities through policies was significant to this objective.

We obtained information from NIH on NIH-hosted or supported large-scale repositories of de-identified, individual-level human genomic information and from CDC on the four CDC repositories containing human genomic data, as well as the data access level—such as whether the data were open, registered, or restricted—of the human genomic data in these repositories.[2] We examined the data security policies or guidelines specific to any NIH- and CDC-hosted repositories of human genomic information—including those from the *All of Us* Research Program, an NIH-supported consortium with a registered-access data repository. In addition, we obtained information from NIH on data management or security violations associated with these repositories, including NIH's description of the incident, and coded the nature of the violation, whether restricted-access data were involved, and whether a penalty was imposed. In coding the data management or security violations, we determined them to be commonly occurring with evidence of three or more of that type of violation. We included all data

[2]Repositories providing only aggregate or summary results, as well as providing small, non-individually-unique amounts of data were not included in the scope of our report.

management or security violations in our analysis to provide complete data on violations of NIH's Genomic Data Sharing Policy.

We conducted this performance audit from February 2024 to April 2025 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient and appropriate evidence to produce a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Appendix II: Comments from the Department of Health and Human Services

DEPARTMENT OF HEALTH & HUMAN SERVICES

OFFICE OF THE SECRETARY

Assistant Secretary for Legislation
Washington, DC 20201

April 21, 2025

Candice N. Wright
Director, Science, Technology Assessment,
and Analytics
U.S. Government Accountability Office
441 G Street NW
Washington, DC 20548

Dear Ms. Wright:

Attached are comments on the U.S. Government Accountability Office's (GAO) report entitled, **"HUMAN GENOMIC DATA: HHS Could Better Track Use of Foreign Testing Entities and Strengthen Oversight of Security Measures"** (GAO-25-107377).

The Department appreciates the opportunity to review this report prior to publication.

Sincerely,

Mitchell Hailstone

Mitchell Hailstone
Acting Assistant Secretary for Legislation
and
Principal Deputy Assistant Secretary

Attachment

GENERAL COMMENTS FROM THE DEPARTMENT OF HEALTH & HUMAN SERVICES ON THE GOVERNMENT ACCOUNTABILITY OFFICE'S DRAFT REPORT - HUMAN GENOMIC DATA: HHS COULD BETTER TRACK USE OF FOREIGN TESTING ENTITIES AND STRENGTHEN OVERSIGHT OF SECURITY MEASURES (GAO-25-107377)

The U.S. Department of Health & Human Services (HHS) appreciates the opportunity from the Government Accountability Office (GAO) to review and comment on this draft report.

GAO Recommendation 1:
HHS ONS should develop and disseminate training and guidance on supply chain risk assessment standards that enable operating divisions to implement effective risk management for genomic data security while maintaining a focus on their core missions.

HHS Response:
HHS concurs with the recommendation.

The Office of National Security (ONS) has been actively involved in training initiatives focused on the Enterprise Supply Chain Risk Management (E-SCRM) program through a variety of platforms. These include collaborative meetings with the Enterprise Risk Management (ERM) Council, intelligence professionals, and key federal agencies such as the Centers for Medicare & Medicaid Services (CMS), the Food and Drug Administration (FDA), the Centers for Disease Control and Prevention (CDC), the National Institutes of Health (NIH), and the Administration for Strategic Preparedness and Response (ASPR), as well as the Office of Acquisitions within the Assistant Secretary for Financial Resources (ASFR). However, ONS has recently temporarily paused any further training efforts until the Office of the Chief Information Officer (OCIO) finalizes the Information and Communications Technology (ICT) framework. This framework, once completed, will be integrated into our training resources to improve an overall comprehension of SCRM and cyber-related Supply Chain Risk Management (SCRM).

In addition to these efforts, ONS is currently collaborating with our communications specialist and the Assistant Secretary for Public Affairs (ASPA) to develop a training video that will be made available to all personnel within the Department of Health and Human Services (HHS). ONS is dedicated to supporting Operating Divisions (OpDivs) and Staff Divisions (StaffDivs) by offering guidance on the establishment of their own SCRM programs and when ONS should be contacted to conduct Foreign Ownership, Control, or Influence (FOCI) reviews such as when their assessments indicate potential foreign concerns.

NIH Overall Comment:

NIH published updated security standards for controlled-access repositories, their Approved Users, and Developers in Guide Notice NOT-OD-24-157, dated July 25, 2024. In April 2024, NIH prohibited data access requests from researchers located in countries of concern. In August and again in November 2024, NIH reminded GAO about the Guide Notice. While GAO's draft report summarizes some aspects of the NIH Genomic Data Sharing Policy and processes of controlled-access repositories, there is no mention of this security update, which represents a significant change in NIH's security stance. NIH recommends that this update be acknowledged in the final report.

GAO Recommendation 2:
The Director of NIH should direct that NIH begin systematically tracking the extent to which intramural and extramural researchers use genetic services provided by entities with ties to countries of concern.

HHS Response:
NIH concurs with GAO's recommendation and considers it open.

The NIH will provide an update to address the recommendation in our 180-day letter response to Congress.

GAO Recommendation 3:
The Director of NIH should require the development and implementation of procedures to proactively and comprehensively monitor researcher compliance with data management and security measures for human genomic data.

HHS Response:
NIH concurs with GAO's recommendation and considers it open.

As described in detail during the engagement, NIH is implementing multiple changes that address GAO's recommendation and the underlying concern regarding countries of concern misusing human genomic data. Specifically, in April 2024, NIH prohibited data access requests from researchers located in countries of concern. In July 2024, NIH issued an update to implementation of the Genomic Data Sharing Policy, NOT-OD-24-157, which updated security standards for controlled-access genomic data repositories and their users, which will provide greater assurance of the security of genomic data and opportunities for monitoring researcher compliance. NIH is considering additional actions to protect research participants' data.

The NIH will provide an update to address the recommendation in our 180-day letter response to Congress.

GAO Recommendation 4:
The Director of CDC should direct CDC to develop and implement procedures, across all its centers that maintain restricted-access repositories with human genomic information, to proactively and comprehensively monitor researcher compliance with data management and security measures.

HHS Response:

CDC concurs with GAO's recommendation and considers it open.

CDC will take steps to develop and implement appropriate procedures.

Appendix III: GAO Contact and Staff Acknowledgments

GAO Contact

Candice N. Wright at WrightC@gao.gov

Staff Acknowledgments

In addition to the contact named above, Joseph Cook (Assistant Director), Michael Steinberg (Analyst in Charge), Sara Shore, Benjamin Schaefer, Emily Quick-Cole, Amy Pereira, Jenny Chanley, Mark Kuykendall, and Joseph Rando made key contributions to this report.