Association of
American Medical Colleges
655 K Street, N.W., Suite 100, Washington, D.C. 20001-2399
T 202 828 0400 F 202 828 1125
www.aamc.org

November 29, 2024

U.S. Department of Justice
National Security Division
175 N Street NE, 12th Floor
Washington, DC 20002

Re: Docket No. NSD 104, Notice of Proposed Rulemaking: National Security Division; Provisions Regarding Access to Americans' Bulk Sensitive Personal Data and Government-Related Data by Countries of Concern (89 FR 86116)

Submitted at www.regulations.gov

The Association of American Medical Colleges (AAMC) offers here comments in response to the Department of Justice (DOJ) notice of proposed rulemaking (NPRM) entitled "Provisions Regarding Access to Americans' Bulk Sensitive Personal Data and Government-Related Data by Countries of Concern." In addition to the opportunity to submit these written comments, we appreciate the outreach by the DOJ to encourage comment, answer questions about the intended meaning and impact of the rule through stakeholder meetings, and provide information about the engagement of other federal agencies in the process of drafting the rule.

The AAMC is a nonprofit association dedicated to improving the health of people everywhere through medical education, health care, medical research, and community collaborations. Its members are all 159 U.S. medical schools accredited by the Liaison Committee on Medical Education; 13 accredited Canadian medical schools; nearly 500 academic health systems and teaching hospitals, including Department of Veterans Affairs medical centers; and more than 70 academic societies. Through these institutions and organizations, the AAMC leads and serves America's medical schools, academic health systems and teaching hospitals, and the millions of individuals across academic medicine, including more than 201,000 full-time faculty members, 97,000 medical students, 158,000 resident physicians, and 60,000 graduate students and postdoctoral researchers in the biomedical sciences. Following a 2022 merger, the Alliance of Academic Health Centers International broadened participation in the AAMC by 70 international academic health centers throughout five regional offices across the globe.

General Comments

The NPRM, in its implementation of Executive Order 14117,[1] proposes a rule that will have a profound and potentially unduly burdensome impact on the biomedical research enterprise and the academic institutions where much of this research is developed. AAMC's members conduct the majority of federally funded research in the nation and substantial levels of research funded through other sources including institutional funds, non-federal foundation grants, and industry sponsorship. Although aspects of the rule may impact other operations at medical schools, academic health systems and teaching hospitals, we focus our comments here on the effect of the proposed rule on biomedical research and suggest ways in which the rule could be more easily implemented without compromising the national security risks that DOJ and the White House seek to address through the Executive Order and accompanying regulations.

In response to the previous advance notice of proposed rulemaking (ANPRM) issued by DOJ,[2] the AAMC submitted comments[3] urging the DOJ that the agency avoid issuing regulations which "negatively impact the scientific progress that results from sharing health and biomedical data." We reiterate them here in response to the NPRM which adopted the same definition and approach. In particular, we expressed concerns with the treatment of genomic data, using the same risk framework as other types of data such as financial data and geolocation data. We cited the broad definition of "human genomic data," [4] which captured genomic data types with enormous variability in identifiability and potential security risk. Some of these data, such as a subset of a gene sequence, pose little to no risk of identifying the person from whom the gene originated. We also noted that data security standards for genomics data that have been developed to protect the identifiability and accessibility of the data are typically not tied to the number of individuals represented in the data, but instead are responsive to the safeguards on access to the data, noting: "accepted scientific practice currently examines a number of technical factors to determine the risk level of sharing genomic data and whether the resulting data should be accessible through open access or controlled access."

---

[1] "Preventing Access to Americans' Bulk Sensitive Personal Data and United States Government-Related Data by Countries of Concern," Executive Order 14117, 89 FR 15421, February 28, 2024. https://www.whitehouse.gov/briefing-room/presidential-actions/2024/02/28/executive-order-on-preventing-access-to-americans-bulk-sensitive-personal-data-and-united-states-government-related-data-by-countries-of-concern/

[2] "Provisions Regarding Access to Americans' Bulk Personal Sensitive Data and Government-Related Data by Countries of Concern," Advance notice of proposed rulemaking. 89 FR 15780. March 8, 2024.

[3] AAMC response to DOJ ANPRM: Provisions Regarding Access to Americans' Bulk Sensitive Personal Data and Government-Related Data by Countries of Concern. Submitted April 19, 2024. https://www.aamc.org/media/75841/download?attachment

[4] "The term *human genomic data* means data representing the nucleic acid sequences that comprise the entire set or a subset of the genetic instructions found in a human cell, including the result or results of an individual's 'genetic test' (as defined in 42 U.S.C. 300gg–91(d)(17)) and any related human genetic sequencing data." 89 FR 15785.

The AAMC is disappointed that after sharing these concerns with the DOJ through written comments and with the Science and Technology Policy Institute, which surveyed organizations and individuals on behalf of the DOJ, the NPRM retains the framework for regulation of human genomic data and uses the same principles that are applied to non-scientific data. We understand that our concerns are shared by other scientific organizations and academic institutions and commend those organizations' comments to the DOJ.

The AAMC has three overarching recommendations, as described in this response:

1. **Ensure that when a research study or collaboration is exempt from Subpart C (Prohibited transactions) through a provision in Subpart E (Exempt transactions), all related research activities, required vendor or employment agreements, and transfers for legal or ethics review of the research are also explicitly exempt, regardless of the source of funding for those specific activities.**

2. **Limit the regulation of human biological data to genomic data and do not include any "other 'omic data" in these regulations, as the NPRM provides no evidence nor indication that the other 'omic data present either a national security risk or a likelihood that individuals could be identified solely through this data.**

3. **Recognize that academic institutions have neither the existing cybersecurity nor investigatory infrastructure of for-profit companies or the federal government, and therefore the burden of compliance will be unduly costly for academic institutions without significant assistance through guidance, tools, and clear mechanisms for engaging in beneficial, well-governed international research collaborations. Failure to take these attributes into consideration in finalizing this rule would result in an overbroad rule that would discourage research collaborations, to the detriment of the United States. International research collaborations advance science, keep the United States globally competitive in research and development, and encourage countries that are not countries of concern to continue engaging with researchers and research institutions in the United States.**

Prohibited transactions and exempt transactions

The proposed rule generally prohibits any "covered data transaction" with a covered person involving human genomic data or biospecimens from 100 or more U.S. persons (see discussions of covered person and bulk data thresholds, below). The provisions that delineate exempt transactions do exempt those "conducted pursuant to a grant, contract, or other agreement entered into with the United States Government" (§202.504(a)). **We agree that in the context of federally funded research these already-regulated activities should not be prohibited.** We note, however, that in conversations with stakeholders the DOJ has implied that to meet this

exemption the transfer itself must be specifically funded by the federal grant. In general, any given academic research program or laboratory at an academic institution is supported by a combination of federal and non-federal grants, and the proposed exemption would still prohibit a significant amount of scientific data exchange that could be subject to the same level of institutional oversight as those research activities specifically included in federal grants.

The exemptions for regulatory approval data (§202.510) and clinical investigations (§202.511) would allow human genomic data transfer for industry-funded clinical trials that are designed to support marketing applications for drugs, devices or biological products, but would not allow similar transfer for epidemiologic, public health, foundational, or preclinical research, none of which are intrinsically more risky than the existing exemptions. **We recommend the incorporation of an exemption to allow Covered Data Transactions in academic research studies regardless of funding source. This exemption could include a requirement for reasonable assurances, contractual obligations, or audit rights as necessary.** We recognize that this goal could also be accomplished through a general license as described in Subpart H and anticipate that the DOJ could move quickly to identify the scenarios for which these general licenses should be issued.

The exemptions at §202.510 and §202.511 are too limited in scope to ensure that clinical investigations, among other research, can be conducted in compliance with the proposed rule. Prior to clinical trial data being submitted to international regulatory authorities for product approval, ethics bodies such as institutional review boards must review the research protocols and sometimes preliminary data prior to research commencing. During the course of a trial, countries monitor the progress and adverse events of research happening locally. Committees such as data safety monitoring boards (DSMBs) may include local experts or authorities or may be based in the country where the research is taking place. Without the ability to transfer certain data in these situations, trials may be halted or prevented from starting long before the existing exemptions would apply. **We urge that the existing exemptions be expand to include required activities prior to regulatory submission for product approval, including ethics review and trial monitoring for adverse events during the course of the trial.**

Obtaining regulatory or ethics approval for international research routinely requires local experts who are sometimes also key collaborators. Experts who know the language, research context, and navigation of local regulatory bodies can be essential to accomplish these exempt activities. **Thus, we recommend that the exemptions for regulatory approval data and clinical investigations be further expanded to include the vendor or employment agreements reasonably required to conduct activities related to obtain regulatory or ethics approval for research in a country of concern. In these exemptions, the regulations could require documentation that the otherwise prohibited or restricted data transactions are necessary to accomplish the exempt activity.**

The NPRM exempts data transactions to the extent that they are required or authorized by: (1) The Pandemic Influenza Preparedness and Response Framework; (2) The Global Influenza Surveillance and Response System; and (3) The Agreement between the Government of the United States of America and the Government of the People's Republic of China on Cooperation in Science and Technology (1979). While we appreciate these exemptions for public health and pandemic preparedness and agree with their inclusion in the rule, we note that the nature of emerging infectious diseases or other global health emergencies could warrant unanticipated and urgent data sharing needs which are not covered by the three frameworks or agreements listed in the proposed rule. **We suggest that §202.507(b) include a mechanism, similar to the issuance of a general license in §202.801, to allow academic researchers to respond rapidly to a global emergency and share genomic data if needed, if the crisis does not fall under the provisions of the three listed agreements. We also ask that the assertion in the preamble, that the definition of human genomic data "does not include non-human data, such as pathogen genetic sequence data, that is derived from or integrated into human genomic data" be incorporated directly into the rule, which will prevent an interpretation that would hamper immediate investigation of pathogens.**

Exclusion from "sensitive personal data"

We were glad to see the categorical exclusion for data that is contained in "unrestricted and open access repositories" from the definition of "sensitive personal data." These are well-known, widely-used platforms for worldwide sharing of genomics data. We support the inclusion of this exclusion in the final rule.

Regulation of "other 'omic" data

The NPRM does not include in the proposed provisions, but asks in the preamble, about the regulation of "other human 'omic data" beyond genomic data, suggesting that the regulations could potentially include "human epigenomic data, glycomic data, lipidomic data, metabolomic data, meta-multiomic data, microbiomic data, phenomic data, proteomic data, and transcriptomic data." The inclusion of these data in the regulations would have a profoundly negative impact on scientific research. As described below, this rule as drafted will require all academic institutions to review *every research project with an international component or collaborator, regardless of whether or not that component or collaborator is located in a country of concern.* Identifying the universe of such research that involves human genomes alone is a substantial effort. Broadening that search to include all researchers who study the downstream data from other biological processes could readily sweep in countless other explorations, the vast majority of whom are not studying "data from U.S. persons." Without any indication of the specific national security risk posed by these other 'omic data, it does not make sense to capture these considerations in this current effort.

The AAMC strongly advises against the inclusion of any other 'omic data in this regulation at this time, finding that the potential burden of compliance and broad chilling effect on biomedical research is not outweighed by any information in the preamble nor in information readily available to AAMC that suggests these data have similar risk profiles, sensitivity, or likelihood of identifiability as genomic data.

Each of these other 'omics are scientifically distinct and have risk profiles and characteristics that vastly differ from each other, as well as from genomic data. Similar to genomic data, the risk profile of these data types cannot be determined in a volume-based manner by the number of individuals represented in the data. As set forth in the Executive Order, any proposed regulations "shall not address transactions to the extent that they involve types of human 'omic data other than human genomic data" before the submission of the report that would assess "the risks and benefits of regulating transactions involving types of human 'omic data other than human genomic data… and [recommend] the extent to which such transactions should be regulated." No information contained in the preamble mentions or identifies any specific risks of these other 'omics; all examples and citations apply only to genomic data. We do agree, however, that these materials and data should be appropriately assessed for their potential risks and the scientific costs of regulating these transactions. **To ensure that the most relevant and current scientific expertise informs any future rulemaking related to potential restrictions or prohibitions on sharing other 'omic data, we recommend the formation of a review panel to provide information and recommendations for future revisions to these regulations through the review mechanism described in the NPRM. We recommend inclusion of scientists from the federal government, as selected by the Director of the NIH, and from academic institutions.**

<u>Ensuring regulations can be implemented by the regulated community</u>

Ensuring regulatory compliance requires clear, specific, and implementable regulations for the academic research community. Lack of clarity around definitions for covered data transactions or requirements for individual institutions to conduct risk analysis or independently identify covered persons will create a regulatory framework that makes institutional compliance extremely difficult. Such complexity will likely lead to institutions taking a maximally restrictive approach to any international collaborations, stifling research progress and increasing undue regulatory burden in conducting research.

Identifying covered persons

The due diligence required to identify all covered persons, as described in §202.211, exceeds the capabilities of most academic institutions, requiring resources and access to information that few, if any, academic institutions will be able to meet. Academic institutions have, however, been focused on addressing concerns about foreign interference in research since a 2018 letter to all NIH grantees, warning about the agency's concerns regarding undue foreign influence on

federally funded research. A government-wide effort led by the White House Office of Science and Technology Policy (OSTP) since 2018 on a comprehensive approach to research security has resulted in academic institutions implementing policies and procedures to identify, assess, and restrict or monitor interactions and collaborations with international persons or entities.

These institutions have created processes to identify and review the primary affiliations of collaborators who reside in other countries, particularly countries of concern. The proposed rule, however, sets forth a definition of covered person that would also require an institution to know, for example, if an individual with access to research data that included any genomic data or biospecimens: resides anywhere in the world (other than the U.S.) and is an employee or contractor of a country of concern; resides anywhere in the world and is an employee or contractor of an entity that is outside of a country of concern *but* is 50 percent or more owned *directly or indirectly* by a country of concern or is chartered in a country of concern. This definition will require that institutions review the contracts, agreements, and employment status every collaborator for each research project in every country in the world, resulting in a daunting, nearly impossible effort to ensure compliance. Further complicating this effort is the fact that there will not be a comprehensive list of covered persons or of those entities chartered in countries of concern with locations in other countries. **The rule or accompanying guidance should provide explicit permission for academic institutions engaged in research to be able to rely on certifications with documentation to determine when an international researcher *outside of a country of concern* is a covered person.**

The AAMC urges the DOJ to recognize in these efforts the existing and emerging federal framework for research security that is currently being finalized by OSTP, with input from across the federal government, including security and intelligence agencies. Collaborating with OSTP and other relevant agencies in order to streamline requirements and ensure their alignment is critical to ensure that institutions are not subject to dual or conflicting regulations and policies. We request that DOJ indicate in any subsequent rulemaking how efforts to safely and securely protect and research data are being coordinated across the whole of government.

We note that the overlay of this proposed regulation on existing research security requirements, cybersecurity developments, and export control regulations does not allow institutions to leverage current efforts and resources to ensure compliance with these new requirements as has been assumed in the NPRM. In many cases, the definitions that do not match current institutional policies and the threshold-based risk assessments that do not align with other risk-based frameworks (those that assess risk by activity instead of by type of data) will make these rules more difficult to implement rather than easier.

Economic impact of compliance, overcompliance, and foregone transactions

The NPRM has calculated the likely impact of foregone transactions that would be prohibited under this rule. What is not accounted for are the *permitted* collaborations that will not be initiated as a result of the requirements. While that is certainly true for permitted, beneficial research with collaborators who reside in countries of concern, it is also the case that a rule that cannot be readily implemented or well understood will also impact international research *that has no connection to any person or entity in a country of concern.* This is an outcome that we must prevent, working together through partnership between the federal government and the research community. As described above, due diligence requirements that cannot be met and exclusions that only cover parts of research projects or certain stages of scientific advancement have the effect of chilling all international research collaborations based on fear of noncompliance.

Other Recommendations

In response to concerns regarding the definition of "human genomic data," the DOJ declined to make any revisions to the scope of the definition, which includes any "subset of the genetic instructions found in a human cell." As we explained in our response to the ANPRM, coming to a consensus on an appropriate threshold for "bulk" data when considering both a fully sequenced human genome and also any part of a gene. In many cases, without linked identifiers, a DNA sequence of a single gene or targeted area of the genome being studied would pose no risk of reidentification of the persons from whom the original samples were collected, regardless of the number of U.S. persons included. These activities would not raise the concerns the rule is attempting to address and should be excluded from Subpart C through a readily-available mechanism. **We recommend that the DOJ establish, through an exemption in the regulation, quickly developed guidance, or general licenses, that the threshold of 100 U.S. persons apply only to fully sequenced entire genomes and that there be a process for demonstrating and certifying that the risk of re-identification of a particular data set is so low that it is exempt from prohibited transactions.**

Conclusion

AAMC agrees with the use of risk-based safeguards when sharing or transferring sensitive data, and strongly supports the Executive Order's assertion that any new policies should preserve the "open, global, interoperable, reliable, and secure flows of data across borders." We also appreciate the document's explicit support for "open scientific data and sample sharing to accelerate research and development through international cooperation and collaboration." We remain concerned, however, that the sweeping prohibitions on biomedical research and the extraordinary cost to academic institutions of ensuring compliance with the rule as drafted will have an immediate and significantly negative impact on the United States' engagement in

international research, *even in institutions that do not conduct or propose any prohibited or restricted transactions with covered persons.*

International collaboration is a key component of the U.S. research enterprise and greatly strengthens both our capacity and ability to advance scientific knowledge. This type of information exchange is particularly critical during quickly moving global health emergencies such as the COVID-19 pandemic. AAMC-member medical schools and teaching hospitals routinely work with sensitive biomedical and health data, conducting and managing research that is increasingly data-driven and often national or international in scope. Academic institutions have already committed to sharing the responsibility for protecting research data from foreign interference, protecting intellectual property from foreign governments, and addressing national security threats related to identifiable research data. The suggestions and concerns raised here are offered to ensure that the United States remains globally competitive, fully engaged in the international research community, and able to meaningfully contribute to the threat of emerging diseases and to global health needs. The AAMC urges the DOJ to deeply consider the concerns of the scientific community as it moves to finalize this regulation to prevent an unintentionally deleterious effect on medical research and scientific progress.

Please feel free to contact me or my colleagues Heather Pierce, JD, MPH, Senior Director of Science Policy and Regulatory Counsel (hpierce@aamc.org) and Anurupa Dev, PhD, Director of Science Policy and Strategy (adev@aamc.org) with any questions about these comments.

Sincerely,

Elena Fuentes-Afflick, MD, MPH
Chief Scientific Officer

cc: David J. Skorton, MD, AAMC President and Chief Executive Officer