COLLEGE of AMERICAN PATHOLOGISTS

March 18, 2026

The Honorable Jayanta "Jay" Bhattacharya, MD, PhD
Director
National Institutes of Health
Department of Health and Human Services

Re: Request for Information on Draft NIH Controlled-Access Data Policy and Proposed Revisions to NIH Genomic Data Sharing Policy (Notice Number: NOT-OD-26-023)

Submitted via Electronic Submission to https://osp.od.nih.gov/comment-form-draft-nih-controlled-access-data-policy-and-proposed-revisions-to-nih-genomic-data-sharing-policy/

Dear Director Bhattacharya,

The College of American Pathologists (CAP) appreciates the opportunity to comment on the Request for Information (RFI) on the *Draft NIH Controlled-Access Data Policy and Proposed Revisions to NIH Genomic Data Sharing Policy*. As the world's largest organization of board-certified pathologists and leading provider of laboratory accreditation and proficiency testing programs, the CAP serves patients, pathologists, and the public by fostering and advocating excellence in the practice of pathology and laboratory medicine worldwide. As physicians specializing in the diagnosis of disease through laboratory methods, pathologists have a long track record of delivering high quality diagnostic services to patients and other physicians.

The NIH Notice NOT-OD-26-023 requests public input on two proposed policy changes aimed at better protecting human research participant data. The first is a new NIH Controlled-Access Data Policy that would establish clear, standardized requirements for which types of human data must be stored and shared through controlled access. Data types covered include genomic, proteomic, transcriptomic, and epigenomic data, along with personal health and financial data, biometric identifiers, geolocation data, and facial/head imaging. Repositories handling this data must meet security standards including identity verification, and access review. The policy is partly driven by national security directives, including Executive Order 14117 and the Consolidated Appropriations Act of 2023, which restricts sharing of biospecimens and data with Countries of Concern.

The second proposal updates the 2014 NIH Genomic Data Sharing (GDS) Policy with the stated goal of reducing redundancy and modernizing its requirements. Key changes include narrowing the policy's scope to human genomic data only, simplifying the "large-scale" threshold to any study with 100 or more participants, standardizing requirements across all NIH Institutes, and requiring data submission to approved repositories within 6 months of generation. If consent has not been obtained, genomic data generated from biospecimens collected after 2015 cannot be shared. The update imposes new consent requirements for sharing genomic data (including requirements for consent from next of kin or applicable legal authority in cases where the individual is deceased, consent from a proxy decision maker in cases where the individual lacks capacity to consent and in

College of American Pathologists
1001 G Street, NW, Suite 425W
Washington, DC 20001
202-354-7100

the case of children, assent from minors with parental permission). It also allows HIPAA Expert Determination for de-identification and expands which institutional bodies can approve data submissions beyond just IRBs. NIH is accepting written public comments through March 18, 2026 via the NIH Office of Science Policy website.

Comments:

1) **Clarification of the Scope and Applicability**:
It is not clear if the policy changes would apply to specific research projects funded by the NIH or to all research conducted in an institution that receives NIH funding. Clarification of this point would be helpful. Furthermore, under Scope and Applicability, the RFI states, "This Policy applies to all NIH-supported research generating human data or deriving data from human data, cell lines, or biospecimens" but then almost immediately states, "This Policy does not apply to NIH research that only involves: Collection and sharing of human cell lines and biospecimens." Does this mean that if a biobank is simply collecting, storing, and distributing human biospecimens but not generating research data from those biospecimens, the biobank would not need to comply with the policy universally for all its specimen collections, but the end users the biobank distributes biospecimens to would for specified research? And if the biobank is distributing de-identified biospecimens and associated personal health data (e.g., diagnosis) the policy does apply because the biobank is now collecting, storing, and distributing personal health information? Most biobanks collect, store, and share some form of health information. Biospecimens are essentially useless without associated clinical data. Also related to applicability, biobanks often provide supportive services (biospecimen and associated clinical data) for NIH-funded investigators, but are not typically conducting research. If biobanks are supporting research, but not necessarily conducting research, does this policy apply?

   In addition, it is not clear how these proposed policy changes would apply to researchers who are storing protected data types in systems that are not considered a controlled access data repository and what the requirements for controlled access would be if they are not sharing with a NIH Controlled Data Access Repository. The RFI points to multiple related policy announcements that include definitions and controlled access requirements. The RFI does not include the definition of a controlled access data repository in its list of definitions. This is included in the related policy announcements referenced in the text. It will be critical for researchers and institutions to better understand what constitutes a controlled access data repository and the requirements for sharing controlled access data as specified in the policy, particularly for an individual researcher whose dataset does not meet the definition of a controlled access data repository. What will the requirements be for sharing types of "omics" data that are not "genomics" data?

2) **Data Types to be Protected**: Several new data types have been added to the list of data that must be protected through controlled access under the policy. These now include epigenomic data, proteomic data, and transcriptomic data. We understand that these data types are identified as types of data covered under the DOJ's Final Rule "Preventing Access to American's Bulk Sensitive Personal Data and United States Government-Related Data by

Countries of Concern or Covered Persons." However, we note that the sharing of these data types may not pose the same level of risk of identifiability and potential harms as sharing genomic data. While some genomic sequence may be inferred from some types of proteomic and transcriptomic data this typically is not sufficient to uniquely identify an individual, hence has generally not been considered "high risk" research. On the other hand, in the case of rare diseases, sharing even small datasets may constitute an increased level of risk. While the intention of extending the NIH policy to additional data types beyond genomic data may be to protect against national security risks, the sharing of such data beyond a certain threshold (>100 persons for genomic data, >1,000 persons for other "omics" data) with Countries of Concern is already prohibited under the DOJ Rule. Consideration should be given to providing more flexibility in the NIH policy for sharing "omics" data based on the potential level of risk.

3) **Expansion of the Definition of Genomic Data**: The RFI includes the following definition of genomic data: "*Data representing the nucleic acid sequences that constitute 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 (28 CFR 202.224)." 42 USC 300gg-91(d)(17) defines a "genetic test" as "an analysis of human DNA, RNA, chromosomes, proteins, or metabolites that detects genotypes, mutations, or chromosomal changes."* Thus, the policy changes being considered would now appear to expand the definition beyond whole genome sequencing/whole exome sequencing and include single gene and somatic (cancer-related) genetic testing. We note that single gene and somatic genetic testing do not have the same identifiability risks to individuals and the same implications for families as whole genomic sequencing. The rationale for this expansion in the definition may be to further align the definitions of the NIH policy with the DOJ Rule, but this may place additional burdens on institutions and negatively impact important research collaborations. The sharing of datasets of certain sizes of "omics" data is already prohibited with Countries of Concern under the DOJ Rule, so it is not clear what this policy would do beyond impeding sharing of "omics" data for important research.

4) **Requirements that Go Beyond the Common Rule**: One of the major problems with the proposed policy is that it imposes requirements that go beyond the Common Rule in several areas. The Common Rule only applies to the use of information or biospecimens in which there is interaction or intervention for the collection of biospecimens or data, or the use of identifiable private information. The Common Rule does not consider genomic or other "omics data" covered by the draft NIH policy inherently identifiable. Rather identifiability is defined as "readily ascertainable by the investigator." Therefore, the Common Rule does not require consent for the use and sharing of genomic or other "omics" data unless there was interaction/intervention with the research participant and/or the identity of the participant is readily ascertainable by the investigator. Furthermore, the Common Rule permits an IRB to waive informed consent for minimal risk research if certain conditions are met. Finally, the Common Rule does not apply to the use of biospecimens or data from deceased individuals. The policy changes proposed in this RFI would apply to genomic data regardless of whether it is identifiable under the Common Rule and to genomic data from deceased, as well as

living individuals, and would require explicit consent for sharing of such data. This policy imposes consent requirements that go beyond the Common Rule and would, in effect, negate recent efforts to harmonize federal requirements (by OHRP and FDA) throughout the research enterprise.

5) **Consent Requirements that May Hinder Important Research**. The proposed policy has consent requirements that would hinder important research in several areas. The policy requires explicit consent for the sharing of genomic data generated from biospecimens or cell lines created or collected after 2015 regardless of whether the data are identifiable or from deceased individuals. As mentioned above, we note that the Common Rule allows the use and sharing of biospecimens for the generation of such data if the conditions for a waiver of consent have been met or the biospecimens and data are not readily ascertainable to the investigator. However, the current and NIH proposed policies do not allow such waivers. Obtaining consent poses significant challenges and barriers for the use of certain types of tissue, as explained further below. Such tissue is a rich source of human biospecimens primed for omics (particularly genomic, epigenomic, and spatial transcriptomic) data generation.

   Types of research which would be hindered by the policy include the following:

   A) **Important Research with Genomic Data Generated from Archival Biospecimens which May Have Been Collected During the Course of Routine Care After 2015** (such as those from pathology archives) for which such research is initially not anticipated and therefore without participant consent. We strongly agree that whenever research is anticipated using human biospecimens and cell lines, that consent be obtained whenever possible. However, sometimes research using archival biospecimens is not anticipated at the time the biospecimens are collected from routine care and therefore, consent is not obtained. Getting consent after the fact may be difficult and sometimes impossible because it may be difficult or impossible to locate patients afterwards. The strict requirements for consent for the sharing of genomics data (or any “omics” data) derived from such biospecimens collected after 2015 would inhibit important research using these samples.

   B) **Important Research Using Data Derived from Pediatric Biospecimens.** If consent is required for sharing of genomic data (or other “omics” data) obtained from pediatric patients when they reach the age of majority, some important research will be impeded. For research involving children, it is important for the ethics committee to have the option to waive the requirement for informed consent for genomic data sharing in certain circumstances. When a minor legally reaches adulthood, consent from the now adult research participant needs to be considered. This also applies to previously collected biospecimens or data that are still being utilized, or if access to medical records is necessary. The requirement for consent, however, can be waived by an IRB under the US federal human subjects regulations (Common Rule and FDA) if the criteria for a waiver of consent is met based on a justified request provided by the investigator. For example, the research participant cannot be found because he/she changed his/her

name, or the parents/custodians do not know his/her address or whereabouts. In addition, a waiver of consent can also be approved by an IRB for the use of remnant/residual biospecimens (those obtained following a medical procedure) or other data included under the proposed policy for research that does not meet the definition of human subjects under the US federal human subjects regulations; (that is, the researcher cannot readily ascertain the identities of the individual from whom the biospecimen is obtained). This approach enhances remnant tissue utilization and associated data via new technologies to facilitate cutting edge research that can ultimately improve the health care of children and expand the use of rare pediatric biospecimens (e.g., samples from pediatric sarcomas, leukemias, and adolescent/young adult cancers) for discovery phase studies. Reconsenting patients who reached the age of majority can also raise important ethical issues. The final policy should provide guidance on the assent requirements under the policy, which should be consistent with current requirements under the Common Rule and FDA.

C) **Important Research Using Data Derived from Biospecimens from Deceased Individuals.** The requirements for consent from next-of-kin (NOK) for genomic data sharing of data from deceased individuals not only goes beyond the Common Rule requirements, but is fraught with practical, ethical, legal, and cultural challenges. For example, finding and contacting family members after an individual's death is often challenging due to outdated contact information, and relatives may be dispersed or estranged. The resources required for finding the NOK may be prohibitive. Even when families can be located, determining who legally and morally qualifies as NOK can be complicated by a number of factors, such as determining who is NOK in the case of blended families, unmarried partners, or cultural differences in family structure. Selection bias may be introduced into datasets by high non-response rates and skew research outcomes.

There are also unresolved ethical issues, such as questions about how well the NOK can represent the autonomy and privacy interests of the deceased potential research participant, and, in the case of genomic data with implications for family members, whose voice should count. Approaching grieving families to obtain a research consent can feel intrusive and/or create anxiety, especially if the families were not expecting such contact. Finally, the regulatory landscape regarding the authority of NOK may vary across jurisdictions and may be ambiguous.

These issues are particularly acute for the sharing of genomic data from pathology archives where the potential research participants may have died long ago. Again, while genomic data sharing from biospecimens collected prior to 2015 are grandfathered in, this would preclude the use of valuable biospecimens for genomic studies in the future for which research was not initially anticipated and therefore consent not obtained prospectively. Decisions about whether consent for genomic data sharing of deceased individuals is best left to the IRB to weigh the risks and benefits when making decisions about institutional certification for genomic data sharing.

Finally, it is not clear if the proposed requirement for informed consent for genomic data sharing from NOK applies if the individual deceases after consent has already been obtained from the individual for genomic data sharing. This issue needs clarification as it would raise additional ethical issues regarding autonomy interests.

D) **Response to Pandemic Outbreaks:** During a pandemic, rapid data sharing is needed. The requirements proposed, particularly for explicit consent without the possibility of a waiver by an IRB, would hinder or even prevent rapid sharing of data needed to address emerging public health crises, such as the recent COVID-19 pandemic.

6) **Institutional Burden and Impact on Certain Populations:** The requirements of the proposed policy would impose new costs and burdens to implement data sharing systems and negatively impact certain populations.

As stated in the RFI, the goal of the proposed policy changes is aimed at better protecting human research participant data. However, regulations related to the collection, use, and sharing of potentially individually identifiable biospecimens and information are already in place, and it is unclear to what extent the additional requirements of the proposed policy improve upon these current rules and regulations. Furthermore, it is anticipated that implementation of this policy will significantly increase the need for regulatory, operational, and administrative oversight of human biospecimens, associated health information, and derived “omics” data with potentially limited impact on improving the protection of human research participant data beyond that already provided by existing HHS and DOJ policies. As such, there seems to be an imbalance between the resources required to implement the proposed policy and the overall benefit to the protection of human research participant data. In essence, the policy changes proposed could be considered an unfunded mandate.

The burdens would be particularly problematic for small institutions, academic institutions, and sites in low- and middle-income countries. These costs and burdens, particularly in low-resource settings, could not only ultimately negatively impact new scientific discoveries and continued innovation, but also negatively impact representation of certain populations in research, exacerbating inequities and limiting the benefits of research and medical care to these populations.

7) **Impact on US Competitiveness in Research:** The requirements proposed by the current policy (particularly for consent that goes beyond the Common Rule) may significantly reduce the number of samples available for research involving the sharing of data covered under the policy, as well as impede certain research as described previously. These barriers have the potential to hinder the US in its competitiveness in the research arena.

College of American Pathologists
1001 G Street, NW, Suite 425W
Washington, DC 20001
202-354-7100

COLLEGE of AMERICAN PATHOLOGISTS

* * * * *

We appreciate the opportunity to submit these comments and would welcome the opportunity to discuss our concerns with the National Institutes of Health. Questions can be directed to Suanna Bruinooge at sbruino@cap.org.

Sincerely,

Qihui "Jim" Zhai, MD, FCAP
President
College of American Pathologists

College of American Pathologists
1001 G Street, NW, Suite 425W
Washington, DC 20001
202-354-7100