Comments are due September 29, 2021

NIH RFI: Developing Consent Language for Future Use of Data and Biospecimens

Authors: Alicia Hofelich Mohr, Katie Barrick, Caitlin Bakker, Shanda Hunt

https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-131.html

Information Requested

NIH is interested in input on 1) the sample consent language, 2) the "points to consider," and 3) any gaps or additional components that should be included. NIH is also interested in input on any hurdles or barriers to the voluntary use of the sample language and "points to consider" by the community. NIH welcomes input from research investigators, institutional review board members, study participants, professional organizations, associations with a focus on research oversight, and other interested members of the public. Respondents are free to address any or all of the information listed below or any other relevant topic for NIH to consider. Respondents should not feel compelled to address all items.

Resource for Stakeholder Input Follows:

Consent for Data and Biospecimen Sharing for Future Use: Points to Consider and Sample Language

I. Introduction:

As a steward of the nation's biomedical research enterprise, NIH is dedicated to ensuring that when data and biospecimens are shared, that it is done ethically and securely, and with respect for the privacy, autonomy, and well-being of research participants and the communities to which they belong. As part of this commitment, NIH is working with stakeholders to identify best practices for developing and

implementing effective consent practices to inform prospective research participants about potential risks and benefits of data and biospecimen sharing for future research. The following resource outlines suggested points to consider when addressing data and biospecimen storage and sharing in consent language and provides supplemental sample language that could be modified as needed when constructing informed consent forms. Of note, the sample language provided below is intended to serve as a helpful resource and is not a substitute for addressing federal, state, local, or tribal requirements that may apply to informed consent. Use of the information provided in this resource, including sample language, is completely voluntary.

II. Instructions for Use:

This document presents points to consider, instructions for use, and optional sample language that is meant to supplement informed consent forms for research studies that include the storage and sharing of data and biospecimens. This resource is neither a linear nor comprehensive consent template. Additionally, the sample language does not address all possible scenarios for which informed consent may be needed for data and biospecimen storage and sharing. The sample language will need to be tailored to institutional and study specific requirements. It is the responsibility of investigators and institutional review boards (IRBs) to determine the appropriate use of the sample language including which components, if any, are relevant to a specific study's informed consent and the most appropriate section to incorporate the sample language within when doing so (e.g., the risks of storage and sharing may be included in the study's informed consent "risk" section or in another appropriate section). Not all of the components will be appropriate for every informed consent form. Investigators should carefully select language appropriate for the study, and IRBs should ensure that the proposed language meets all applicable regulatory and policy requirements, including federal, state, local, and tribal requirements.

Use of this sample language is completely voluntary. This language is being provided as a resource for the research community and there are no requirements that any portion of the language be used in an informed consent form for an NIH-supported or -conducted study.

This resource consistently refers to "data and biospecimens" as a means to capture all identifiable information and biospecimens that research participants may contribute as part of a research study. "Data and biospecimens" includes information collected from, or about a research participant during the course of a primary study (e.g., surveys, medical images, electronic health records, wearable device information) as well as human material (e.g., blood, tissue, urine, extracted DNA).

Some sample language includes embedded instructions to fill in specific information pertaining to the research study. These embedded instructions are identified in **[bold, bracketed text]** and will need to be replaced after study-specific language is inserted or removed entirely based on the instructions provided.

III. General Points to Consider:

- Data and biospecimens may involve distinct storage and/or sharing procedures. Some protocols may require separate consent language to inform how data versus biospecimens are stored and shared.
- Those responsible for study conduct and oversight are encouraged to consider the reading level of the entire informed consent form, with the goal of creating understandable language that conveys the necessary information. The sample language in this resource was crafted to ensure an appropriate reading level (with a goal of ~8th grade reading level or below). Additional resources on evaluating readability can be found from the National Cancer Institute (NCI).

- Studies that involve a category of participants who are considered vulnerable to coercion or undue influence, such as children, prisoners, individuals with impaired decision-making capacity, or economically or educationally disadvantaged persons, or research with pregnant women, fetuses or neonates may require additional considerations regarding the storage and sharing of data and biospecimens. Those responsible for study conduct and oversight are encouraged to revise the sample language to reflect these considerations. We strongly encourage consultation with the appropriate contacts to determine and take into consideration the applicable regulations, policies, and laws relevant to studies involving these populations, including assent for participants under 18, prior to storage and sharing of data and biospecimens.
- Some cultural/donor/sovereign groups may have preferences or requirements regarding how data and biospecimens are handled, including the disposition of biospecimens. For example, sovereign Tribal Nations may have laws/regulations/policies governing research that may impact the storage and sharing of data and biospecimens. We strongly encourage consultation with the appropriate contacts to determine applicable regulations, policies, and cultural preferences or tribal laws that will need to be taken into consideration prior to storage and sharing of data and biospecimens.
- Additional considerations may be applicable for research studies that include the storage and sharing of genomic data. We recommend that those responsible for study conduct and oversight review community standards, such as NIH resources provided by the National Human Genome Research Institute (NHGRI) on informed consent and the NIH Genomic Data Sharing Policy.
- If the future use of data and biospecimens will be limited, this information should be specified in the consent language.
- As technology advances for coding and deidentifying data and biospecimens, consider the implications for privacy and confidentiality and adjust language as appropriate.

IV. Sample Language Components:

Component 1: Introduction - Description

Considerations for those responsible for study conduct and oversight: The Introduction-Description component is meant to provide prospective research participants with an introduction to, and description of the storage and sharing of data and biospecimens in the study.

- If participants may be re-contacted to collect new or replacement data or biospecimens, include language to address re-contacting.
- Those responsible for study conduct and oversight will need to consider the appropriate timeframe for data and biospecimen storage based on their study and anticipated uses. For some, the appropriate timeframe may be indefinite, while others may have a clear, limited timeframe.

Instructions for those responsible for study conduct and oversight: See sample language below for the Introduction-Description component. If using this sample language, include the first three paragraphs then choose either Option #1 or Option #2. Replace embedded instructions identified in **[bold, bracketed text]** with specific information pertaining to the study and remove **[Option #1 and #2 text]**.

Sample Language:

This study is collecting data and biospecimens from you. We would like to make your data and biospecimens available for other research studies that may be done in the future. The research may be about similar diseases or conditions to this study. However, research could also be about unrelated diseases, conditions, or other aspects of health. These studies may be done by researchers at other institutions, including commercial entities. Our goal is to make more research possible to learn about health and disease.

Your data and biospecimens will be stored **[indicate the name of the institution where they will be stored, including any biobanks to be utilized]**. *We plan to keep your data and biospecimens for* **[indicate time frame or "indefinitely," or until "used completely," etc.]**.

Your data and biospecimens may be shared with investigators around the world. However, access to the data and biospecimens is controlled by **[indicate which entity has control].** *To use your data and biospecimens, researchers must get approval and they must agree not to try to identify you.*

[Option #1: If the data/biospecimens are coded and can be linked back to the participant]

We will protect the confidentiality of your information to the extent possible. Your name and other identifying information will not be on any data and biospecimens you provide. The data and biospecimens will have a code that links to your identifying information. The code key will be kept in a locked location separate from your information. The code key can only be accessed by people who have permission.

[Option #2: If the data and biospecimens are completely delinked from identifiers and cannot be linked back to the participant]

Your name and identifying information will not be on any data and biospecimens you provide. Investigators cannot link your identifying information to the data and biospecimens.

Component 2: Voluntary Participation

Considerations: The Voluntary Participation component informs prospective research participants about the voluntary nature of data and biospecimen storage and sharing.

- In general, participants should be given the option to agree to, or opt out of, having their data and biospecimens stored and shared for future research. Providing options for participants to agree to, or opt out of, having their data and biospecimens stored and shared is particularly important in studies that offer the prospect of direct benefit to the participant. Mandating agreement to storage and sharing may be considered coercive if the participant does not want to agree to sharing of data and biospecimens but feels compelled to agree anyway in order to join a possibly beneficial clinical trial. If the research protocol offers no prospect of direct benefit, then it may be reasonable for storage and sharing not to be optional.

- If the protocol is a repository protocol with the sole intent of collecting data and/or biospecimens for future use, no opt out mechanism is necessary.

Instructions: Choose either Option #1 or Option #2. Remove **[Option#1 and #2 text].**

Sample Language:

[Option #1: When sharing of data and biospecimens will be optional (e.g., for studies that have potential benefit)]

It is your choice whether or not to let researchers share your data and biospecimens for research in the future. If you say "yes," you can change your mind later, but your data and biospecimens might still be used if they have already been shared. If you say "no," you can still fully participate in this study. Please initial next to your choice:

_____YES, *use my data and biospecimens in other research studies*

_____NO, do NOT use *my data and biospecimens in other research studies*

[Option #2: When sharing of data and biospecimens will not be optional (e.g., for studies where sharing is integral to the purpose of the study)]

Participating in this study means you agree to share your data and biospecimens. You can change your mind later, but researchers may still use your data and biospecimens that have already been shared. If you do not want your data and biospecimens used for other projects, you should not participate in this study.

Component 3: Discontinuation/Withdrawal

Considerations: The Discontinuation/Withdrawal component describes what will happen if the participant changes their mind about storage and sharing.

Instructions: Adjust language as necessary.

Sample Language:

You can change your mind about sharing your data and biospecimens at any time. If you change your mind, please contact the study team to let us know. We will not share your data and biospecimens going forward. We will do our best to retrieve all your data and biospecimens that have already been shared, but it may not be possible. For example, if some research with your data and biospecimens has already been done, the information from that research may still be used. We will not know which data and biospecimens are yours if the identifying information was removed. Also, if the data and biospecimens have been shared already with other researchers, it might not be possible to get them back.

Component 4: Risks & Benefits

General Considerations: The Risks & Benefits component describes the reasonably foreseeable risks/discomforts related to storage and sharing of data and biospecimens, and any benefits related to storage and sharing of data and biospecimens that prospective participants may receive.

Considerations - Risks: If identifying information (e.g., key to the code) will remain with the data and biospecimens during storage and sharing, include language that addresses the additional measures designed to safeguard participants' privacy (e.g., access controls).

- Ensure that the safeguards listed are consistent with language addressing the storage and sharing of data and biospecimens in the introduction.
- Adjust language if there is a specific risk associated with loss of privacy due to storage and sharing, such as stigma or the ability to obtain certain types of insurance.

Instructions: Adjust language as needed. Remove [**Risks**] and [**Benefits**] unless needed as a section heading.

Sample Language:

> [**Risks**] *When we share your data and biospecimens, there is a small risk that people may get access to it who are not supposed to. We will protect your data and biospecimens as much as possible during storage and when they are shared. However, there is a small chance your identity could be discovered.*
>
> [**Benefits**] *You will not receive any direct benefit from sharing your data and biospecimens. However, sharing your data and biospecimens may contribute to research that helps others in the future.*

Component 5: Commercial Application

Considerations: The Commercial Application component informs prospective participants about whether their data and biospecimens may contribute to products with commercial value. If research participants will receive any payments related to commercial or product development, adjust language in the last sentence to reflect this.

Instructions: Adjust language as needed.

Sample Language:

> *The use of your data and biospecimens may lead to new tests, drugs, devices, or other products or services with commercial value. These products or services could be patented and licensed. There are no plans to provide any payment to you should this occur.*

Submitting a Response

Comments should be submitted electronically by September 29, 2021, using the form at https://osp.od.nih.gov/rfi-comment-informed-consent-sharing/.

You may provide comments to one or all of the topics in the comment boxes. Comments received will be posted at https://osp.od.nih.gov/clinical-research/informed-consent/ without change after NIH has reviewed all of the comments received. Please do not include any proprietary, classified, confidential, or sensitive information in your response.

This Request for Information (RFI) is for planning purposes only and should not be construed as a policy, solicitation for applications, or as an obligation on the part of the Government to provide support for any ideas identified in response to it. NIH may use information gathered by this RFI to inform development or modification of websites, policies and practices, processes and procedures, and supporting documentation.

Our response draft

Domain of research most important to you or your organization (e.g. cognitive neuroscience, infectious epidemiology)

Academic institutional repository & research support group

1. Utility and useability of this resource (limit: 8000 characters)

- Overall, we are very glad to see this guidance coming from NIH and we recognize that consent language for data sharing and biospecimens is a challenging area, for researchers, funders, and regulatory bodies.

2. Gaps or additional components that should be included (limit: 8000 characters)

- This guidance seems out of sync with the data management and sharing guidance taking place in 2023. If NIH is mandating the sharing of data, they should take a stronger stance on encouraging researchers to INFORM their participants that their data or biospecimens will be shared. Should at the very least strongly encourage. Should also mention that several repositories that take in NIH data do require non-restrictive consent language.
    - The Final NIH Policy for Data Management and Sharing defines data sharing as "[t]he act of making scientific data available for use by others (e.g., the larger research community, institutions, the broader public), for example, via an established repository."
- There should also be a general notice that the language suggested here is designed for researchers who plan to share data in a limited access repository or in some other access-mediated way, not for those who plan to make their data publicly available. We would like to see language options for public or more open-access repositories as options in this document.

    - Our perspective is that of an open access repository that ingests human participant data. The guidance recommended in this proposal does not support sharing in our repository.

3. Specific language proposed in the informed consent sample language (limit: 8000 characters)

- IRBs need to be trained on data sharing and best practices. IRB responses to questions about data sharing in the absence of informed data sharing consent may not always align with the best ethical practices.
- Encourage giving specific examples of who explicitly should serve as "contacts" for questions of vulnerable populations and cultural/donor/sovereign groups. Unclear if this means community groups, IRB/national offices, or other institutional groups (data repositories, HIPCO offices, etc).
    - Also clarify when these consultations should be occurring - before data collection? Only when storing/sharing? Both?
- Need to define what "limited" sharing might mean. Does this mean in terms of restricted access to data by researchers? That researchers might approve of data sharing for only certain uses? Maintained for a certain amount of time?
- Referring to "approval" in terms of who has access to the data: much of the proposed language here refers to a managed, restricted access repository, rather than a public-access repository. Institutional repositories, used frequently by academia, are often completely open access. This will impact the language that is used and what participants need to be told.
- Need to expand on and be more specific about the "technology advances and de-identification" to include consideration of the fact technologies for re-identification are also improving, and that de-identification methods that "worked" at one time with one dataset may not be effective in the future and as more datasets are made available. Give suggestions or language on how researchers can make participants aware of the fact that things may not be perfectly "de-identified"
- Be specific about who has "permission" to access the "code key" to the data - research team and those who have been granted access to the data.
- Regarding “If the research protocol offers no prospect of direct benefit, then it may be reasonable for storage and sharing not to be optional,” does not need to be all or nothing. Maybe this is a matter of wording: “If the research protocol offers a direct benefit, then data storage and sharing should be optional.”
    - Is this sentence even necessary?
    - We encourage consent language that explicitly states data will be shared.
- Regarding “If the protocol is a repository protocol with the sole intent of collecting data and/or biospecimens for future use, no opt out mechanism is necessary,”
    - It’s not 100% clear if “opt out” means opting out of the study or opting out of data sharing.
    - If it means opting out of data sharing, repository protocol could be better defined - for example, an institutional repository often has the sole intent of collecting data

for future use, but that doesn't mean participants shouldn't be able to opt out of sharing their data. In fact, it's perhaps even riskier given the open nature of most institutional repositories.
    - Repository protocol is overall a vague term - we are unsure what, exactly, is meant. An operational definition of repository protocol, including reference to existing standards (such as ISO OAIS model, TDR, or Coretrustseal), should be provided in order to clarify.

4. Hurdles or barriers to wider use of this resource by the community (limit: 8000 characters)

- See "Gaps or additional components" above.

5. Other considerations relevant to this resource (limit: 8000 characters)

- It would also be helpful to provide guidance for researchers on how to choose the type of access (or repository, if applicable) appropriate for their data and how to determine the best language to inform participants of the way their data will be shared. We would also like to see explicit mention of when these decisions about data sharing should be made, and encourage researchers to make them BEFORE submitting IRB protocols.
    - Suggest explicitly pointing to existing materials such as the Supplemental Information to the NIH Policy for Data Management and Sharing: Selecting a Repository for Data Resulting from NIH-Supported Research (https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-016.html)
- Consider recommending to research teams that they educate research participants about data sharing beyond the consent form, highlighting benefits for science, openness, and transparency that come with researchers making the data from their studies available to others.
    - UMN data sharing information template as an example (https://docs.google.com/document/d/19GuL5TJCDx3DiU59kWmiTh_E64O1pSZ40uCQvyBW0Ns/edit?usp=sharing)
    - This may also have alignment with NIH's recently released strategic plan, which emphasizes the importance of data sharing, transparency, public engagement, and encouraging participation of underrepresented groups in clinical trials. Public education and outreach to participants are mechanisms by which those goals can be operationalized within a single study.