In response to a Request for Information (RFI) from the National Institutes of Health (NIH) regarding Strategies for NIH Data Management, Sharing, and Citation (NOT-OD-17-015), FASEB submitted the following comments. The RFI presented eight topics on which NIH sought input from stakeholders, with responses on each topic limited to 250 words. FASEB’s comments were submitted electronically via web form at http://osp.od.nih.gov/content/nih-request-information-strategies-nih-data-management-sharing-and-citation on December 7, 2016.

SECTION I. Data Sharing Strategy Development

1. The highest-priority types of data to be shared and value in sharing such data

Improving data management and increasing data access can create new scientific opportunities. The Federation of American Societies for Experimental Biology (FASEB) agrees with NIH that many factors must be considered when determining what data should be shared and under what conditions. We also applaud NIH for seeking to identify the types of data that are of the highest priority to share. For this purpose, FASEB recommends the potential utility of a dataset to be used as a major criterion. NIH’s approach to data sharing should recognize that access to some datasets may not be worth the cost of sharing and long-term preservation.

(Please see our attached Statement for additional information.)

2. The length of time these data should be made available for secondary research purposes, the appropriate means for maintaining and sustaining such data, and the long-term resource implications

The diversity of data types, research areas, and resources available make it challenging to identify data accessibility strategies that are practical and relevant for all fields of research supported by NIH. Therefore, the Federation of American Societies for Experimental Biology (FASEB) encourages NIH to employ flexible approaches that allow investigators and NIH to establish reasonable expectations for a particular research project. Data management plans (DMPs) are an important tool for clarifying these expectations while still providing flexibility at the research project level.

(Please see our attached Statement for additional information.)

3. Barriers (and burdens or costs) to data stewardship and sharing, and mechanisms to overcome these barriers

Making datasets accessible – including the skilled human labor necessary to prepare and maintain data, technological infrastructure, and continued development of databases and effective search platforms – is costly. Expansion of requirements for data sharing will require commensurate financial and staff support...
from research sponsors. The Federation of American Societies for Experimental Biology (FASEB) affirms that NIH should, at a minimum, provide sufficient support to fully comply with all applicable data management and access requirements as part of a project’s funding.

Additional resource needs and barriers exist; to identify common challenges, NIH can examine data management plans (DMPs) in aggregate and continue seeking feedback from the scientific community.

(Please see our attached Statement for additional information.)

4. Any other relevant issues respondents recognize as important for NIH to consider

The Federation of American Societies for Experimental Biology (FASEB) encourages NIH to ensure that investigators are empowered to carry out quality data management. Poor data management practices can render a potentially valuable dataset useless. Investigators may require specialized software and other tools, training, topical or data type-based repositories, and assistance locating relevant resources. NIH should continue addressing these needs across the biological research enterprise.

In addition, community-based data standards can enhance the utility and interoperability of datasets. NIH can serve as a leader and convene of stakeholders to develop these standards as needed. Scientific societies can assist by identifying experts, providing thoughtful feedback, and disseminating proposed standards.

Finally, FASEB notes that some datasets have little value for reuse or a short “shelf-life.” Requirements to share and preserve such data could create inefficiencies in research funding and resource distribution.

(Please see our attached Statement for additional information.)

SECTION II. Inclusion of Data and Software Citation in NIH Research Performance Progress Reports and Grant Applications

1. The impact of increased reporting of data and software sharing in RPPRs and competing grant applications to enrich reporting of productivity of research projects and to incentivize data sharing

Data and software citation in RPPRs and competing grant applications can serve as a limited source of professional recognition for investigators that share these research products.

(Please see our attached Statement for additional information.)

2. Important features of technical guidance for data and software citation in reports to NIH

[No comments were submitted for this topic]
3. Additional routes by which NIH might strengthen and incentivize data and software sharing beyond reporting them in RPPRs and Competitive Grant Renewals applications

To incentivize data sharing, the Federation of American Societies for Experimental Biology (FASEB) recommends that NIH take actions that maximize the value of shared datasets and assure professional recognition for making them accessible. Investigators have little incentive to share data if they do not think other researchers will be able to find, use, or cite it. FASEB recognizes NIH’s efforts in these areas and encourages NIH to build upon this important work.

NIH’s support for the development and provision of resources and services that enable data sharing should be sustained. This includes providing funding for databases and the development of training modules as well as NIH staff assistance to identify available resources and repositories. Continued effort is also needed to promote productive data sharing, from building better systems for data discovery to facilitating the development of community-based data standards.

As noted in this request for information, data citation can encourage sharing by ensuring that credit is attributed to the investigators responsible for generating the data. There are further strategies that NIH can employ to advance citation practices. For example, all NIH databases and data catalogs could provide standardized, exportable citation information for datasets. These export tools make it easier for investigators to accurately cite data and provide a visible reminder to do so.

(Please see our attached Statement for additional information.)

4. Any other relevant issues respondents recognize as important for NIH to consider

Integrated community-based solutions are needed to enhance data management and access. The Federation of American Societies for Experimental Biology (FASEB) encourages NIH to work with the research community and other stakeholders to develop and refine its approaches to data sharing. Research sponsors, investigators, institutions, and scientific journal can all contribute to and benefit from advancing data management and access.

Coordination with stakeholders should also extend to other biological research sponsors. Many challenges and needs are cross-cutting. NIH can harmonize policies for data management plans (DMPs) with other federal agencies as well as develop unified portal systems for data discovery. NIH and other sponsors should also conduct continuous assessment of data sharing policies and requirements to ensure that they do not delay the adoption of improved practices or new technologies.

(Please see our attached Statement for additional information.)

Attachment: FASEB Statement on Data Management and Access (March 1, 2016)