Via Electronic Submission

January 19, 2017

Francis S. Collins, MD, PhD
Director
National Institutes of Health
9000 Rockville Pike
Bethesda, Maryland 20892

Subject: NIH Request for Information (RFI): Strategies for NIH Data Management, Sharing, and Citation (Notice Number: NOT-OD-17-015)

Dear Dr. Collins:

The American Society of Clinical Oncology (ASCO) appreciates the opportunity to provide comments and information to help inform the National Institutes of Health’s (NIH’s) plans for data management and data sharing strategies. ASCO represents more than 41,000 physicians and other professionals who care for people with cancer and conduct research to improve cancer treatment.

General Comments

ASCO strongly supports the NIH efforts to develop a framework and strategies for data sharing and data management, which are essential for expediting the translation of research results and the dissemination of best practices to improve public health. ASCO agrees with NIH’s principle that “the results of federally-funded scientific research are made available to and are useful for the general public, industry, and the community”.1 Further, effective data sharing relies upon appropriate identification, adoption, and crediting of good data management and sharing practices, and the principles to make data “FAIR” (Findable, Accessible, Interoperable, and Reusable).2

In addition to these principles, we note two additional factors we strongly urge the NIH to consider in the development of any data management and sharing strategies: 1) additional funding is critical for effective accessibility and utilization of data to occur and

1 https://grants.nih.gov/policy/sharing.htm
2 http://www.nature.com/articles/sdata201618
2) NIH should reconcile various data sharing policies that have emerged from several offices within the agency, and harmonize its policy (to the greatest extent possible) with other requirements researchers face. These requirements include results reporting on ClinicalTrials.gov and journal requirements for data submission as well as emerging journal requirements for data sharing. These points will also be detailed below in our response to the specifics within this Request for Information (RFI).

Our comments below are provided in the context of goals and strategies we have considered in the development of two ASCO data sharing platforms, ASCO’s rapid learning system, CancerLinQ, LLC (CLQ), and the Targeted Agent and Profiling Utilization Registry (TAPUR), a prospective clinical trial. Both initiatives have highlighted the importance of common data elements and structured data collection to enable searching for and sharing information and integrating data across datasets.

The development of our rapid learning health care system, CLQ, will allow clinicians to analyze aggregated, real-world cancer clinical data from electronic health records (EHR). Clinical trials like ASCO’s TAPUR Study include multiple therapeutic options and operate across all cancer types. The TAPUR Study leverages rapid advances in tumor genomic sequencing and broadening availability of such testing to facilitate participation by cancer patients with all cancer types. The goals of these initiatives are to drive better cancer diagnosis and treatment.

**Information Requested**

**Section 1. Data Sharing Strategy Development**

In this section the NIH notes that many factors must be considered when determining what, when, and how data should be managed and shared. These factors include, for example, the purpose for sharing, supporting data re-use and reproducibility, maturity of the science, the uniqueness of the data, and ethical considerations.

The CLQ and TAPUR platforms are examples of the valuable types of data to be shared because they represent the experiences of the diverse patient populations treated in real world clinical practice. The value in sharing such data is the ability to learn from every patient, with the expectation that this will accelerate progress against cancer and provide patients and physicians more comprehensive information to guide decisions about cancer prognosis and treatment. The sharing of data in a timely manner, once completed studies are reviewed and accepted for publication, provides researchers and clinicians the ability to deliver better treatments to the right patients at the right time.

General policies and aspects of ASCO’s data sharing platforms include:

- Providing diagnostic and clinical data (i.e. patient demographics, diagnostic tests, treatments administered, clinical outcomes, adverse events, and patient-reported outcomes) to enable translation of basic science findings into clinical outcomes. This is particularly important in the era of precision medicine.
- Standardizing the data elements to ensure they can be used in a broad range of applications.

- Offering thorough documentation and guidance to ensure that investigators not only understand the data but also the initial purpose for its collection, data sources, validation processes, methods and tools that were used for collection and analysis, and recommended analytic methods to employ.

- Fostering a process that encourages and enables subsequent researchers who use the data to share with the original researcher information about data curation, data mapping, and methods that were used to work with the data to ensure that others can learn from any new methods or enhancements.

- Requiring good data stewardship, including a commitment to privacy, strong data security and transparency about the uses of the data.

In addition to these principles, there are a few complex aspects that can cause barriers to data sharing including the need for standardization of data. Below, ASCO has previously shared a few of these key provisions with the agency in 2015, in response to its NIH Request for Public Comments on the Draft NIH Policy on Dissemination of NIH-Funded Clinical Trial Information (NOT-OD-15-019):

1. Strongly recommend NIH to take the lead and work with stakeholders in the community to develop standardized common data elements and encourage researchers to share data in the agreed-upon format. ASCO is particularly supportive of requiring more standardized common data elements in electronic health records and other databases because this will make it easier for users of the data to search for information and integrate it with other datasets. Databases are only valuable to the extent that they are high-quality, well organized, and are standardized across entries.

2. In clinical trials, strongly recommend NIH follow the FDA standard requiring information about the attribution of adverse events to promote consistency in reporting for IND studies, and also use this standard for non-IND studies. The attribution of adverse events is an example of the broad category of data annotation which is essential for effective data sharing. Cancer is a complex and deadly disease with disease-related adverse events. In addition, trials often combine the investigational agent with standard of care treatments. Both of these facts make it difficult, at times, to determine whether an adverse event is caused by the investigational agent, the standard of care agent(s), or the underlying cancer. Thus, it is important that the information on adverse events accurately reflect what information the researchers are reporting (i.e., whether the record includes all adverse events or only adverse events that the researchers believe are attributable to the study drug or intervention).

3. Strongly recommend NIH to consider in the development of its strategies that it is impractical to ask researchers to comply with different data sharing rules, thus guidance is needed on how to coordinate the various approaches regardless of sponsor.

---

4. Strongly recommend NIH to allow researchers to request funding as an allowable budget item in NIH grants to support resources required for ClinicalTrials.gov reporting requirements; maintaining, organizing, and cleaning of data sets and to provide access to researchers who request the data; and the acquiring and analysis of data sets needed for research. This would ease the implementation of the NIH policy by providing researchers with the resources to meet these requirements. It would also align the new policy with the 2015 Institute of Medicine report, *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risks*, which recommended that sponsors and funders “provide funding to investigators for sharing of clinical trial data as a line item in grants and contracts.”

Section 2. Inclusion of Data and Software Citation in NIH Research Performance Progress Reports (RPPR) and Grant Applications

NIH grantees are required to report “other products of the research” in their Research Progress Reports (RPPR) and grant applications including data, databases, and software, in section C5a of their annual RPPR submission. ASCO agrees that more thorough reporting of data and software products in the RPPR and in Competitive Grant Renewal applications may strengthen documentation of productivity and may also identify projects and investigators who most effectively share data and software.

The NIH also requests in this section additional routes by which NIH might strengthen and incentivize data and software sharing beyond the benefits that result from reporting in RPPRs and Competitive Grant Renewals applications. As noted above, incentives should include additional resources to enable effective data and software sharing. This includes not only the preparation of data to be shared, but also the hosting of data and the associated long-term resource implications. We strongly urge the NIH to identify strategies to improve researchers’ compliance with other existing and developing rules governing data sharing. Guidance is needed on how to coordinate various approaches and harmonize policies on data sharing that would otherwise overlap for stakeholders who conduct research projects that must adhere to requirements for ClinicalTrials.gov, various institutes and centers within the NIH (i.e. NIH Data Sharing Policy, NIH Public Access Policy, and the NIH Genomic Data Sharing Policy), trial sponsors, the U.S. Food and Drug Administration, and journal editors.

As another example of overlapping developing policies within the area of data sharing, the recent NIH RFI *Including Preprints and Interim Research Products in NIH Applications and Reports* (NOT-OD-17-006) sought input on whether preprints and other interim research products should be included in NIH applications and reports, and how investigators could report them. In particular, we believe that studies reporting clinical data that have a potential impact on the lives of patients should continue to be shared according to the current standards used within research. Currently, authors are allowed to submit and present abstracts (i.e. oral or poster presentations) of their research in open, scientific meetings. These are considered preliminary data that are not generally viewed by physicians as sufficiently mature to

---


justify a change in clinical practice, but extremely important for continuing education and knowledge on the advancements in science. Further, the current standard is a proven approach for the dissemination of preliminary results, for establishing research priorities, to provide evidence of independence and productivity, to gain feedback from peers, correct errors, and ensure the protection of patient privacy.

In conclusion, ASCO urges the NIH to develop a comprehensive data management and sharing framework that reflects the principles of responsible sharing of data in a timely way and that maximizes the benefits to research, care, and the general public. A fundamental principle of all NIH-funded research is that the results must be disseminated in order to contribute to the general body of scientific knowledge and, ultimately, to the public health. We agree and believe NIH awardees should continue to be expected to make their data and accomplishments of their activities available to the research community and to the public at large for the greater good. We encourage the NIH to continue working towards a broad and coordinated framework that will enable researchers and clinicians to learn from and extend the work of others, to readily share those insights, and to translate research results into clinical advances for patients.

We look forward to working with the NIH toward these important goals.

Thank you for the opportunity to provide comments on this RFI on Strategies for NIH Data Management, Sharing, and Citation. Please contact Shimere Williams Sherwood (Shimere.Sherwood@asco.org) at ASCO with any further questions.

Sincerely,

Daniel F. Hayes, MD, FASCO, FACP
President, American Society of Clinical Oncology