

Senate HELP Committee NIH RFI Comment

This document is submitted by the Massachusetts Health Data Consortium (MHDC) and its Data Governance Collaborative (DGC) in response to Senate Help Committee NIH RFI posted by Senator Bill Cassidy on the committee website on September 29, 2023 and found here:

https://www.help.senate.gov/imo/media/doc/nih_reform_rfi.pdf

About MHDC

Founded in 1978, MHDC, a not-for-profit corporation, convenes the Massachusetts's health information community in advancing multi-stakeholder health data collaborations. MHDC's members include payers, providers, industry associations, state and federal agencies, technology and services companies, and consumers. The Consortium is the oldest organization of its kind in the country.

MHDC provides a variety of services to its members including educational and networking opportunities, analytics services on both the administrative and clinical side (Spotlight), and data governance and standardization efforts for both clinical and administrative data (the Data Governance Collaborative/DGC and the New England Healthcare Exchange Network, respectively).

About DGC

The DGC is a collaboration between payer and provider organizations convened to discuss, design, and implement data sharing and interoperability among payers, providers, patients/members, and other interested parties who need health data. It is a one stop interoperability resource. The DGC primarily focuses on three areas:

1. Collaboration: Development of common understanding of and specifications for data standards, exchange mechanisms, and what it means to participate in the modern health IT ecosystem
2. Education: helping members understand their regulatory obligations, the data and exchange standards they're expected to use, and modern technology and related processes
3. Innovation: Identification and development of projects and services needed to make modern health data practices and exchange a reality

General Comments

This section general comments that cross multiple questions asked in this RFI or that do not have specific questions.

Goal of Increasing the Speed of Science

Our Data Governance Collaborative questions whether increasing the speed of science is the correct goal. While there are special circumstances – such as dealing with a pandemic – where speed is most certainly essential, under more normal circumstances correct and, in some cases, comprehensive seem better than fast. Certainly, getting to correct more efficiently would be a laudable goal and improving processes is always worthwhile, but speed in of itself may be more harmful than beneficial.

Questions about current state

The participants of the DGC with research experience were surprised to see certain questions in this RFI. They felt some of the general process areas covered and questions being asked are well baked into the landscape and working fairly well at this point. They particularly called out some of the questions around public-private partnership as covering ground that seems to be working well. While there is always room for improvement in

any area and with any process, they question whether this is the best use of limited time and resources.

Data standardization and interoperability mindful of privacy

We will comment further under specific questions asked below, but one major area for improvement identified was in promoting/adopting data standardization and better interoperability within the constraints of good privacy practices. This topic is one of our focuses generally, and we identified multiple areas where we feel it would be helpful within the areas covered by this RFI.

Sharing data across multiple research sites working on the same project

This topic did not appear to come up in the specific questions asked so we'll address it here. NIH should consider using standardized data and data sharing methods across multiple research centers working on the same project. We noticed that a recent announcement about a new maternal health project involving multiple sites had to name a data coordination entity to handle data collection, standardization, and analysis (<https://www.nih.gov/news-events/news-releases/nih-establishes-maternal-health-research-centers-excellence>). Members of our Data Governance Collaborative with experience in this type of role indicate that currently it usually involves a lot of manual work with all sorts of different files and data formats coming in to be combined and analyzed. If the data collection and sharing happened using a standard format and standard API calls the need for this sort of entity would be greatly reduced, or at least they would only be needed for analysis and not for collation and conversion tasks.

Response to Specific Questions – Increasing the Pace of Science

This section will list specific questions asked in the Increasing the Pace of Science section of the RFI and provide our responses to them.

What specific policies or systems would better expedite open sharing of NIH-funded data and analyses?

Our Data Governance Collaborative believes there are several improvements that could be made in this area including:

1. Making clinical data from NIH or NIH funded projects available in relevant FHIR formats
2. Developing NIH-sponsored FHIR implementation guides for research data, research findings, and other relevant types of data
3. Require use of common data formats and interoperability/data sharing requirements as part of all NIH research grants
4. Ensure that all data originating from or used in NIH research includes adequate provenance so anyone using the data later understands its origins and how the data flows across organizations and systems
5. Make sure that information about peer reviews (when appropriate) accompanies all research findings in a standard, easy to understand format
6. Have a mechanism for ensuring basic clinical data (standard lab work, radiology, exam results, etc) is exchanged back to the primary care provider and payer (as applicable) so that patients are not asked to repeat these items if needed for their normal medical care, to meet quality measures goals, or for other clinical or administrative purposes and to ensure the patient retains access to these results via payer Patient Access APIs or standard provider access mechanisms. Since the patient identity should be known by the collector/administrator of relevant data even in double blind studies, this level of interoperability should be possible.

How do common NIH funding mechanisms support or discourage

transformational science? How could these funding mechanisms be improved to prioritize IT?

Using a general definition of IT to mean application of technology to assist tasks and processes, we believe that requiring or facilitating the exchange of data should be a priority. This could range from requiring some or all of the suggestions above to using standardized FHIR exchange mechanisms for acquiring clinical data about research subjects (with appropriate consents) to more fully and efficiently acquiring medical history and information about active patient concerns, diagnoses, and treatments to requiring use of a standards-based interoperable way to search external clinical records for potential research subjects (while maintaining proper privacy and consent controls) and gather contact information for outreach programs.

What are the benefits and challenges associated with the current approach to negotiating facilities and administrative (F&A, or “indirect”) costs? How could this approach be changed to maximize the proportion of federal funds going toward direct research expenses? How, if at all, does the current process for negotiating indirect cost rates advantage or disadvantage certain institutions over others?

We are not familiar with current approaches to indirect costs, but we feel that the funding structure should make it as easy as possible to use a portion of the funds to modernize and improve data and IT practices without it harming the evaluation of proposals or being counted against the researchers in any way, especially if grants include requirements or preferences for more standardization and interoperability.

Do you see opportunities to improve the current process for structuring peer review committees? What attributes does NIH tend to prioritize when selecting both chartered and ad hoc reviewers?

We believe that efforts should be made to promote diversity on peer review committees and, in particular, to include people being served by the project on any review (If research is specific to patients with a particular disease, condition, disability, or symptom some effort should be made to include patients in that group who are qualified as full reviewers if possible).

On the subject of peer reviews, we believe there is benefit in giving peer reviewers more access to the raw data used in studies when possible. To facilitate this, we believe the same data formats and FHIR implementation guides recommended for acquisition and sharing of relevant clinical data in the performance of the research should also be used to send raw data to peer reviewers. Having a separate standardized way to record and exchange data about the review itself would also be beneficial.

What role do institutions not affiliated with major research universities, such as other types of academic medical centers or community hospitals, currently play in the NIH ecosystem? How could these types of facilities be more effectively leveraged as research partners?

Leveraging the existing data and interoperability practices of these organizations would be a helpful step. Many of these smaller provider organizations may not have the budget for or as many mandates requiring the use of the more modern data standards and interoperability mechanisms discussed elsewhere in this RFI. Providing them with some of the resources to build that ability would be a boon not just to research but to standard operations within their healthcare services more generally. Until that happens, use of HIEs or other less direct and less widely standardized mechanisms for data exchange may be appropriate (but our Data Governance Collaborative believes this should be considered a temporary workaround and not a long term solution).

Response to Specific Questions – Organizing for Success

This section will list specific questions asked in the Organizing for Success section of the RFI and provide our responses to them.

Does NIH’s current organ- and disease-based structure effectively facilitate the conduct of research? If yes, how? If no, what alternative structure would be more effective in your view? What barriers prevent Congress or the administration from implementing this structure, aside from NIH’s statutory authorization and appropriations?

Having separate institutes means having separate data sharing between NIH and those organizations that supply or use its data. At a minimum, consistent rules for exactly how data should be shared with and by NIH would help. Another option might be having a central data services area that manages the core interoperability between all of NIH and other organizations with proper consent/routing management in place for further distribution within the institutes.

Either way, it would be extremely helpful to have central resources at NIH that set up connectivity/data for individual projects in a consistent way so once someone works with one NIH project they just need new endpoints to connect to another project.

What role could novel technologies, such as artificial intelligence and machine learning, play in protecting the privacy of research participants’ data or inadvertently making this data more vulnerable? What models or capabilities exist to strengthen privacy protections, while improving the timely dissemination of research findings and underlying data?

We believe that more granular consent models are needed to properly protect patient data privacy, especially for uses that fall outside of the HIPAA allowed exchange purposes. Setting up mechanisms that allow patients finer-grained control and enforcing their selections not just in the original data location but throughout the entire lifetime of the data wherever it ends up is an essential part of maintaining privacy and earning patient trust.

Disallowing the use of advertising, trackers, or other privacy-impinging technology on any NIH project would also be a nice step in the right direction. These types of underlying secondary technology, often used without the knowledge or consent of patients, are everywhere and once used the data they collect proliferates and travels to multiple unexpected and unapproved destinations. Similarly, advertising within telehealth platforms or automated check-in processes can have the same effect. See <https://themarkup.org/pixel-hunt/2022/06/16/facebook-is-receiving-sensitive-medical-information-from-hospital-websites> and <https://www.statnews.com/2023/04/07/medical-data-privacy-phreesia/> for two examples.

Having clear and easy to understand privacy and data use policies that outline exactly which identifiable data will be shared across research locations, with peer review, or elsewhere should also be prioritized and some effort should be made to ensure that patients have read them – perhaps having someone summarize the ways data will be shared during an initial intake visit that’s part of a study and thereafter on request might make sense. Providing an overview of how data is de-identified before wider sharing to assure people that it cannot be readily reattached to them might also be helpful.

Members of our Data Governance Collaborative pointed out that there may be cohort-specific sensitivities around data use and privacy. Being sensitive to these is important and is another reason why having diverse staffing that matches the constituencies being served by the research as much as possible is important.

As for the role of AI, it is unclear that AI has a specific role in this area at this time, except perhaps as a monitoring tool to act as a first pass check that policies are being followed.

What are the biggest ethical challenges facing the biomedical research

community today? How is, or is not, NIH currently working to address these issues? What more could NIH do to lead in this space?

Ensuring health equity and providing adequate transparency into the applicability of data/research findings are two of the main ethical challenges. Actively seeking out diverse clinical trial participants is a start, but the same type of diversity is needed for study design, activity, and review. Even more important is being clear about the makeup of trial participants and the characteristics of the population something is tested against. If AI is used to facilitate any of this, clarity around the ground truth data and training models of the underlying algorithms is essential to understand the likely biases inherent in its use.

What specific policy recommendations do you have to improve the transparency of NIH's work, including its accountability to the American people and Congress? Are you aware of any specific mechanisms that have effectively achieved this goal for other federal agencies, including outside of the Department of Health and Human Services (HHS)?

Having frequent webinars, blog posts, email outreach, industry articles, and other public-facing dissemination of information would be helpful. Requiring short quarterly updates of research progress in some type of public venue with perhaps a longer review of findings once finalized might be in order.

Participation in industry events not just in the various fields of research and clinical interest areas but also those covering health equity and/or data standardization and interoperability would help expand the reach of research opportunities and of results.

NIH works with the FHIR accelerator Vulcan, but Vulcan has few public events and does not widely participate in outside events such as WEDI conferences, ONC workshops and conferences, etc. – having both Vulcan and NIH participation in these types of events would greatly increase visibility around health IT in research and encourage more people to work on how best to collect and disseminate research data and findings.

CMS sponsors its own FHIR Connectathon every year that's free to the public and combines CMS, ONC, HL7, and other updates with hands-on collaborative work on various interoperability projects; NIH could consider doing something similar with a research focus and also look to participate in the CMS event.

It would also be helpful to collate related work or ideas across the different institutes so interested parties don't have to try to follow each institutes work independently. Offering a central email list for updates on research opportunities, new rules for research, tips on best practices, data and IT considerations, and more would be helpful in this area too.