



January 27, 2017

Patricia Flatley Brennan, RN, PhD,
Director
National Library of Medicine
National Institutes of Health

Submitted electronically at: NLMStrategicPlan@nih.gov

Re: Notice Number: NOT-LM-17-002, Request for Information (RFI): Strategic Plan for the National Library of Medicine, National Institutes of Health

Dear Dr. Brennan:

The American Medical Informatics Association (AMIA) appreciates the opportunity to submit comments regarding the National Library of Medicine's (NLM) Strategic Plan Request for Information (RFI). AMIA is the professional home for more than 5,400 informatics professionals, representing researchers, front-line clinicians, public health experts, and educators who bring meaning to data, manage information and generate new knowledge across the health and research enterprise.

The NLM is viewed by AMIA, as well as numerous other stakeholders, to be an indispensable and critical component of the NIH. The research it funds, the training it provides, and the infrastructure, tools and resources it makes publicly available is foundational to biomedical informatics and broadly applicable to domain-specific research. Given AMIA's diverse membership, we developed recommendations across all RFI Planning Themes, highlighting areas where consensus was reached.

Specifically, AMIA members envision a future where NLM leads the medical and research enterprise in demonstrating how real-world and “big” health data can be leveraged to develop new concepts of human disease, design novel therapies, and train future clinicians and researchers. Various recommendations discuss the emerging capacity to collect, analyze and apply real-world and big data to improve patient, population and public health. We foresee an opportunity for NLM to take the lead in several aspects of this emerging paradigm, including:

- Providing curated databases of real-world data for use by students, faculty and researchers, initially focusing on EHR data;
- Prioritizing research on the “basic and translational science” of data standards, as well as emerging technologies meant to handle real-world and big data in health;

January 27, 2017

- Focusing future data types and collections relevant to real-world challenges, and available to emerging technologies; and
- Developing technologies, security safeguards, and regulatory models necessary for increasing the ease of sharing and aggregating health data.

In addition, AMIA notes the important role NLM plays in helping various stakeholders apply informatics tools and health IT resources to their unique settings – be it research, clinical or public health. **Specifically, we outline recommendations for how NLM could make sharable and comparable implementation resources developed across the NIH publicly available.** Such resources would help translate and scale evidence-based practices, as not just clinical decision support, but as “intervention decision support.” Information on best practices for implementation of data standards and specific tools developed by various NIH-initiatives would provide tangible value at NLM.

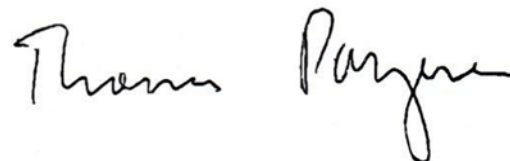
Finally, AMIA envisions a future where patients are first-order participants in their care and in research. **NLM should facilitate this future by supporting development of EHR systems that empower patients through direct access to and interpretation of their health data.** One of AMIA’s Policy Principles is patient empowerment. We believe that policies, programs, research and care delivery should seek to empower patients through access to, and control of, their personal health information. NLM should support research and policies that enable patients to access and transmit all data contained in their EHR, rather than a limited or pre-defined set of data, to improve availability of data for care delivery, biomedical discovery, and in support of patients’ own health and wellness.

Below we outline our recommendations in more detail, and we address NLM’s specific questions related to this RFI. Should you have any questions or require additional information, please contact AMIA Vice President for Public Policy Jeffery Smith at jsmith@amia.org or (301) 657-1291 ext. 113. We, again, thank NLM for the opportunity to comment and look forward to continued dialogue.

Sincerely,



Douglas B. Fridsma, MD, PhD, FACP,
FACMI
President and CEO
AMIA



Thomas H. Payne, MD, FACP, FACMI
AMIA Board Chair
Medical Director, IT Services, UW Medicine
University of Washington

Enclosed: Detailed AMIA Recommendations and Comments to NLM Questions

January 27, 2017

Role of NLM in advancing data science, open science, and biomedical informatics

Identify what you consider an audacious goal that will advance data science, open science and biomedical informatics; a challenge that may be daunting but would represent a huge leap forward were it to be achieved. Include input on the barriers and benefits of the goal.

Given the proliferation of data, and the ways in which we can use those data, across the health ecosystem, AMIA recommends NLM provide curated databases of real-world data, especially EHR data, for use by students, faculty and researchers.

Increasingly, health and healthcare data display the characteristics – in terms of volume, velocity, veracity, variety, value – of big data. While the potential to leverage this data for patient outcomes and discovery looms large, experience to-date has resulted in a “signal-to-noise” ratio that is more likely to cause harm than help. Analytics tools have evolved significantly in the last several years to manage big data, yet students, faculty and researchers do not have open, routine access to these kinds of databases to learn, test and discover.

These challenges are made more difficult by outmoded curricula that do not instill basic informatics literacy and design thinking across general audiences. Both basic informatics literacy and advance applied informatics training is needed to bridge disparate parties who are knowledgeable about a clinical or research domain, yet do not understand what is possible with available data and tools.¹ Design thinking provides a way for unstated, unmet clinical needs to be identified, which would lead to better questions being asked – and a better solution set being developed.²

In order to address this major challenge, NLM should provide de-identified, curated big data that mimics real-world data from EHRs. Subsequently, we recommend NLM look to provide similar data representing biobanks, public health data, and other sources regularly encountered in our digital health ecosystem. These databases should be technology and platform agnostic, and they should include both “clean” datasets, as well as datasets that mimic clean datasets with established “errors” or common variabilities. Such an approach would enable broad cohorts of students, faculty and researchers to use these curated datasets for unique purposes, including for future clinicians interested in leveraging EHR data for quality improvement, and data scientists looking to evaluate results or gaps in real-world data through software and algorithm development.

While we do not anticipate that data donors would be difficult to identify, various barriers of data ownership, curation and maintenance would be of immediate concern. Subsequently, funding and workforce costs associated with sustainability of such databases would be important aspects to determine.

¹ Perlin, J., Baker, D., Brailer, D., Fridsma, D., et al. “Information Technology Interoperability and Use for Better Care and Evidence: A Vital Direction for Health and Health Care.” *National Academy of Medicine*. Discussion Paper, September 19, 2016. Available <http://bit.ly/2d2yFsB>

² Rosen, P., Ku, B., Shah, A., “Making Design Thinking a Part of Medical Education,” *NEJM Catalyst*. 30 June 2016. Available <http://bit.ly/2ju1XGS>

January 27, 2017

AMIA recommends NLM leverage its position as a cross-cutting authority within NIH, as well as a trusted source to faculty, practitioners and researchers outside the NIH, to incentivize open science through data sharing and research rigor. Further, we recommend NLM prioritize multi-disciplinary training programs.

AMIA members note that NLM has an opportunity to ensure its products and services are more broadly applicable to biomedicine because it is not focused on a singular disease. This quality gives NLM the capacity to leverage what other Centers and Institutes do, by abstracting resources that can be generalized and shared. This same characteristic of NLM naturally encourages multi-disciplinary, team-based training, which should be embraced through funding opportunities and curriculum development beyond PhDs.

Given the rapidity of growth and interest in the fields of data science, open science and bioinformatics, NLM should allow for more frequent opportunities to apply for funding and it should more frequently seek ways to partner with other NIH Centers & Institutes on joint funding opportunities. Additionally, NLM should convene non-profit partners, such as IMLS, MLA and AMIA to create bridges between informatics, librarianship, and health professions to fund training programs at the associate's degree (e.g. community college health professions), bachelor's degree, and Master's degree levels. These training programs should leverage role-based informatics competencies.

NLM should also leverage its position to help the NIH address systemic inhibitors to open science. We lack incentives, or have perverse incentives, when it comes to data sharing and academic promotion related to development of useful datasets and software. Important work to finalize 42 CFR Part 11 for depositing data to ClinicalTrials.gov should be supported by similar requirements for research funded by NLM.³ Similarly, NLM should credit scholarship that develops useful datasets and software as a catalyst to overcome systemic disadvantages for such aspects of research in academia and other non-profit research entities.

Most important thing the NLM does in data science, open science, and biomedical informatics

NLM's informatics research and training support is critical to data science, open science and biomedical informatics. AMIA recommends NLM develop ways to enhance knowledge transfer for continued maintenance of these products and services.

Within the NIH, the NLM is the only entity that supports basic informatics research and training, independent of domain. Research, tools, services and skill development funded by NLM has broad applicability across other Centers and Institutes at NIH and contributes research funded by partner agencies AHRQ and PCORI. This focus on informatics research, regardless of domain focus, must remain a central hallmark of NLM.

³ 42 CFR Part 11. Clinical Trials Registration and Results Information Submission; Final Rule.

January 27, 2017

Beyond research support and training, NLM provides indispensable information about, and access to, many standard terminologies / vocabularies needed to harmonize data collected and standardized at the point of capture. NLM provides critical infrastructure for effectively leveraging existing data via the development of tools, programs, and information to enable semantic interoperability.

Resource maintenance is a vital component of NLM. Given the vast number of NLM officials who develop, publish, and refine tools, there is an emerging challenge to make sure the value of those tools is not lost once their authors retire. In order to continue service offerings of tools with high utilization, NLM should emphasize the value of tools that work and the need to keep them maintained. Additionally, we recommend NLM should continue supporting training of medical informatics professionals and researchers, as well as post-master's training that is done of health sciences librarians.

In addition to providing stand-alone resources and tools, NLM provides resources (e.g. the UMLS) that capture the connections and relationships between concepts included in different biomedical terminologies/vocabularies. This facilitates crucial interoperability between healthcare (and research) computer systems, and should continue to be an area NLM excels.^{4,5}

Research areas that are most critical for NLM to conduct or support data science, open science and biomedical informatics

AMIA recommends NLM focus research on the “basic and translational science” of data standards, as well as emerging technologies meant to handle real-world and big data in health.

Data standards are foundational to computation and achieving interoperability. Emphasis on the basic science of data standards is needed, including an examination of how biomedical data standards are (or are not) fit for purpose in health and healthcare. Part of this work should focus on development of granular data standards and information models in clinical and research settings and an acknowledgement that document-centric approaches to interoperability are insufficient. In addition, we recommend NLM support research in how standards are developed and maintained versus how they're implemented, and close the gap between obtaining implementation experience and utilizing the experience to improve standards. AMIA members note that standards are challenged by pragmatic issues related to use in a clinical setting. NLM should partner with clinicians and other stakeholders on particular uses cases to ensure standards are functional and address pragmatic issues.

⁴ Kim TY, Hardiker N, Coenen A Inter-terminology mapping of nursing problems. J Biomed Inform. 2014 Jun;49:213-20. doi: 10.1016/j.jbi.2014.03.001. Epub 2014 Mar 13.

⁵ Adamusiak T1, Shimoyama N, Shimoyama M. Next generation phenotyping using the unified medical language system. JMIR Med Inform. 2014 Mar 18;2(1):e5. doi: 10.2196/medinform.3172.

January 27, 2017

Additionally, NLM should focus research on machine learning and artificial intelligence (AI) as a means to leverage big data, real-world data, and continuous streams data. Such research could also help advance currently resource-intensive methods / processes, such as screening and data abstraction for systemic reviews.

Research into the delta between real-world conditions and controlled environments should be another priority area for NLM to support data science, open science and biomedical informatics. This could include a focus on clinical trial and evaluation outcomes, as well as clinical guideline development versus real-world experience. These efforts could leverage data from EHRs, PHRs, as well as online patient communities, among other sources. It could also focus on consumer health informatics, human-centered data science (intersection of informatics and librarianship) and a broader examination of data issues in global health, e.g. social determinants of health.

New data types or data collections anticipated over the next 10 years to advance data science, open science and biomedical informatics

AMIA recommends NLM focus future data types and collections relevant to real-world challenges, and available to emerging technologies.

As noted above, AMIA recommends NLM endeavor to host curated databases of real-world data, initially focusing on EHR data. NLM should render these and subsequent repositories as reusable reference sets to facilitate machine learning, natural language processing and AI algorithm training.

In addition, we note that various data types will be important to the future of the NLM. Most notably, we anticipate that geospatial, temporal, imaging, environmental, and social determinants data will be essential to understanding health. As required to support Public Health 3.0, A Call to Action to Create a 21st Century Public Health Infrastructure, a “Timely, reliable, granular-level (i.e., sub-county), and actionable data should be made accessible to communities throughout the country, and clear metrics to document success in public health practice should be developed in order to guide, focus, and assess the impact of prevention initiatives, including those targeting the social determinants of health and enhancing equity.”⁶ We recommend that NLM support the collaboration of public and private sectors working to enable more real-time and geographically granular data to be shared, linked, and synthesized to inform action while protecting data security and individual privacy. This includes developing a core set of metrics that encompasses health care and public health, particularly the social determinants of health, environmental outcomes, and health disparities.

⁶ Public Health 3.0, A Call to Action to Create a 21st Century PH Infrastructure Office of the Assistant Secretary for Health, U. S Department of Health and Human Services, <https://www.healthypeople.gov/sites/default/files/Public-Health-3.0-White-Paper.pdf>

January 27, 2017

Other comments, suggestions, or considerations, for advancing data science, open science and biomedical informatics, keeping in mind that the aim is to build the NLM of the future.

AMIA recommends NLM consider ways to ensure that ethical challenges are considered alongside development of new technology, data sources and research.

Increasingly large amounts of data, shared among various stakeholders, brings great promise to solve complex problems. However, this emerging era of big data also brings with it various ethical challenges. NLM should ensure that rich ethics curricula support a growing and diverse workforce – from programmers and data analysts, to clinicians, researchers and administrators. NLM could also consider ways to identify best practices in addressing ethical challenges across various domains, such as genomics and public health. NLM should accomplish these aims by convening stakeholders to serve in an advisory capacity on ethics challenges related to emerging technology, methods and data sharing paradigms.

January 27, 2017

Role of NLM in advancing biomedical discovery and translational science

Identify what you consider an audacious goal that will advance biomedical discovery and translational science; a challenge that may be daunting but would represent a huge leap forward were it to be achieved. Include input on the barriers and benefits of the goal.

AMIA recommends NLM make sharable and comparable implementation resources developed across the NIH publicly available.

Implementation science, or “systems science,” supports fundamental research leading to a theoretical foundation for design and systems engineering that can be integrated and abstracted to develop explanatory models in a general, domain-independent fashion. NLM could work with partner Centers and Institutes, such as NCATS, to make health IT and health informatics application implementation information publicly available. Information on best practices for implementation of data standards and specific tools developed by various NIH-initiatives would provide tangible value at NLM.

AMIA members note translation should not only consider development of new methods, technologies and knowledge to move discovery into practice, but it should also focus on identifying unmet clinical need. This is another area in which NLM could devote resources to design thinking with the purpose of understanding unstated, unmet clinical needs.

Most important thing the NLM does to advance biomedical discovery and translational science

NLM promotes the use of controlled terminologies, uniform data elements and value sets.

Development, maintenance and public access to terminology is a critical service. NLM should continue to promote the use of controlled terminologies, uniform data elements and value sets, and expand such services to other federal agencies, including NIH, CDC, FDA, AHRQ, HRSA, and CMS.

Research areas that are most critical for NLM to conduct or support to advance biomedical discovery and translational science

AMIA recommends NLM focus research on translation of evidence-based practices, and promote availability of negative results to guide implementation.

Research is needed on how to translate and scale evidence-based practices. Intervention Decision Support is needed to guide implementation among disparate groups. For example, our nursing members report that fall prevention guidelines are interpreted and implemented without consultation of major risk scales, which adds to documentation burdens without improving patient

January 27, 2017

care. We need best practices on how to synthesize, streamline, implement and maintain such information and monitor associated outcomes.

Additionally, we support collection and dissemination of negative results, not just positive results through repositories and knowledge bases. Clinical trials represent a significant opportunity, as randomized controlled trials are seen as the gold standard, when real-world use of pharmaceuticals occurs outside controlled environments. Without acknowledgement of the differences in test versus real-world populations, we are constrained by how representative the research is of the real-world. Understanding negative results is also important for machine learning.

January 27, 2017

Role of NLM in supporting the public's health: clinical systems, public health systems and services, and personal health

Identify what you consider an audacious goal in supporting the public's health: clinical systems, public health systems and services, and personal health

AMIA recommends NLM develop resources that (1) contribute to a better experience for clinicians; (2) a more complete understanding of the real-world patient – both inside and outside the clinical setting – and (3) bring transparency to clinical care for patients.

One of the fundamental conceptual advantages of using EHRs remains largely unrealized in terms of data entry mechanisms. AMIA members note that asking clinicians to use contemporary EHRs is akin to asking a baseball player to swing the bat, hit the ball and document what they've done in great detail at the same time. In short, documentation continues to be unpractical. We believe R&D efforts into basic data capture modalities (ambient speech recognition, etc.) are necessary.

Another shortcoming of contemporary EHRs is their inability to collect outcomes information. We know the chief complaint, and we know the intervention, but we do not know whether the intervention actually worked. We have no strong incentives for capturing this information, no good data entry mechanisms, no standards for classifying this information (for example see WHO's thoughts on International Classification of Functioning, Disability and Health (ICF): <http://www.who.int/classifications/BuisnessPlan.pdf?ua=1> Section 2.4). NLM could identify standards gaps and initiate development of ways to collect such data.

Finally, NLM could support development of EHR systems that empower patients through direct access to and interpretation of their health data, including costs, bringing transparency to clinical care. One of AMIA's Policy Principles is patient empowerment.⁷ We believe that policies, programs, research and care delivery should seek to empower patients through access to, and control of, their personal health information. Further we support efforts that enable patients to access and transmit all data contained in their electronic health record, rather than a limited or pre-defined set of data, to improve availability of data for care delivery,^{8,9} biomedical discovery,¹⁰ and in support of patients' own health and wellness. NLM could support research and innovation meant to help patients make sense of their data and leverage their health data for improved care and research.

⁷ AMIA Policy Principles & Positions, Patient Empowerment, available: <https://www.amia.org/public-policy/policy-priorities>

⁸ Klein D., Fix ., et al. (2015). Use of the Blue Button Online Tool for Sharing Health Information: Qualitative Interviews With Patients and Providers. *Journal of Medical Internet Research*. 2015 Aug; 17(8): e199.

⁹ Mohsen, M., Aziz, H. (2015). The Blue Button Project: Engaging Patients in Healthcare by a Click of a Button *Perspectives in Health Information Management*. 2015 Spring; 12(Spring); 1d.

¹⁰ Chisholm, R., Denny J., et al. (2015) Opportunities and Challenges Related to the use of Electronic Health Records Data for Research. *National Institutes of Health Precision Medicine Workshop (Invited White Paper)*. 2015 Feb.

January 27, 2017

Relative to public health, NLM can play critical roles in terms of using technology to leverage public health in the following ways:

- To provide standards (including standard terminologies, uniform data elements and value sets, and standards for public health systems) for public health (including data capture and reporting);
- To provide new approaches and tools to automate / streamline the public health reporting processes;
- To develop new tools and applications to bring in (local and national) instantaneously public health data into point of care dynamically;
- To collaborate with CDC to promote the adoption and implementation of standards in public health.

The standards will improve the reusability and sharability of captured data for public health purpose. This is particularly critical in aggregating data and conducting advanced analysis at population level. There are large overlap between data captured in EHR and public health reporting data or some large surveys (e.g., National Health Interview Survey) in public health. If new tools or applications can capture the right data elements and their values automatically within EHR, then EHR data will have broader secondary usage.

AMIA members also note a need to enhance identification of non-clinical factors to supplement the clinical information above to link population health to environmental, financial, educational and other areas affecting health and well-being. Additional measures to support public and population health include working with other HHS “components to utilize opportunities such as Healthy People 2030, NCVHS’s population health subcommittee, the Evidence-Based Policymaking Commission, and the census to elevate metrics related to social determinant to be leading health indicators, to define community-level indicators that address the social determinants of health, and to explore models to leverage administrative data.”

Research areas that are most critical for NLM to conduct in supporting the public’s health: clinical systems, public health systems and services, and personal health

AMIA recommends NLM support more application-focused research for clinicians, including data capture and knowledge representation in an API-ready ecosystem. In addition, NLM should consider ways to leverage emerging technologies, such as AI, to augment tooling around controlled terminologies.

AMIA recommends NLM support more application-focused research projects. For example, an examination of how information can be customized to provide more relevant, more accurate and more concise information to busy clinicians. Meanwhile the provided information should be able to

January 27, 2017

guide the clinical operations more precisely. The impacts of such research projects are more tangible and straightforward, and should include process and outcome measures.

As stated above, more research on data capture in clinical settings, including data on usage of health IT applications, patients features, treatment processes and patients outcomes. These data will set the foundation for refining clinical treatment procedures, uncovering the possible reasons and their associated factors for different outcomes for patients who have similar conditions and went through similar treatment procedures.

NLM should incorporate machine learning/AI methodology into tooling around controlled medical terminologies. For example, NLM could improve RxNorm's "approximate match" functionality with deep learning algorithms for better text matching. The NLM could expand on the example of RxNorm and drive similar efforts for all key clinical standards that underlie current interoperability efforts (ICD, CPT, LOINC, selected subsets of SNOMED). The goal should be to make it easier to map to these standards and maintain the mappings over time (easy to use APIs are key; good AI underpinning matching functions is necessary).

A deep investigation about (1) how existing standards terminologies, data elements, value sets can be applied in public health data capture and reporting can be a critical starting point; (2) what other standards are needed in terms of public health data capture and reporting can provide guidance for future development.

New data types or data collections anticipated over the next 10 years in supporting the public's health: clinical systems, public health systems and services, and personal health.

NLM should consider ways to collect platform-agnostic decision rules to support clinical decision support (CDS) systems and enable more research on the safety / efficacy of CDS systems in practice.

AMIA members note that CDS is a critical way that informatics can help improve care. NLM should consider a repository of CDS modules linked with value sets maintained in the VSAC as a means to enhance quality and facilitate development of usable CDS for multiple platforms. This would enable a plug-and-play functionality to "hook up" external CDS (which maintains the knowledge and updates such knowledge) with EHRs and provides actionable recommendations. We stress that such CDS needs to focus on real-world scenarios, influenced by real-world data, and with an emphasis on implementable.

January 27, 2017

Role of NLM in building collections to support discovery and health in the 21st century

Identify what you consider an audacious goal in building collections to support discovery and health in the 21st century

Reiterating the need for curated sets of big data, reflective of real-world data, AMIA recommends NLM engage with clinicians to donate de-identified data, similar to how patients have such an option through the All of Us Research Program.

An audacious goal would be to make data collected at the point of care available widely for secondary purposes, such as development of cloud-based decision support modalities, operations improvement, quality measures development and reference implementations, and research.

An analogy could be drawn to increasing calls for requirements to share research data along with a publication. One tactic could be to enable care providers to release de-identified patient data into the custody of NLM in order to make such data available for research. Such an approach could mitigate the challenges and risks involved with developing various one-off common data models and data sharing networks (OMOP, mini-sentinel, PCORnet, SHRINE, etc.).