





DEPARTMENT OF HEALTH AND HUMAN SERVICES  
 NATIONAL INSTITUTES OF HEALTH  
 National Library of Medicine (NLM)

FY 2022 Budget Table of Contents	Page No.
Director’s Overview.....	3
IC Fact Sheet.....	7
Major Changes in Budget Request .....	9
Budget Mechanism Table .....	10
Appropriations Language.....	11
Summary of Changes .....	12
Budget Graphs .....	13
Organization Chart.....	14
Budget Authority by Activity .....	15
Justification of Budget Request .....	16
Program Descriptions .....	16
Appropriations History .....	26
Authorizing Legislation .....	27
Amounts Available for Obligation.....	28
Budget Authority by Object Class .....	29
Salaries and Expenses .....	30
Detail of Full-Time Equivalent Employment (FTE) .....	31
Detail of Positions.....	32





## Director's Overview

### **Mission**

The National Library of Medicine (NLM) is a leader in biomedical and health data science research and the world's largest biomedical library. Our research and information services support scientific discovery, health care delivery, and public health decision making. We pioneer new ways to make biomedical data and information more accessible, build tools for better data management and personal health, and help create a more diverse and data-skilled workforce. Our work enables researchers, clinicians, and the public to use a wealth of biomedical data to improve health. There is not a biomedical discovery, public health advance, or clinical care action in the past 30 years that has not benefited from NLM resources.



Patricia Flatley Brennan,  
R.N., Ph.D.

As 1 of the 27 Institutes and Centers of the National Institutes of Health (NIH), NLM actively contributes to and enhances the discovery and outreach efforts of the NIH. We conduct and support research in computational biology and computational health sciences, and we gather and disseminate biomedical and health data and scientific information resources to enable discovery and health. Our Congressional authorization, as relevant now as it was 65 years ago, charges NLM to acquire, organize, preserve, publish, and make available biomedical and health information to support research and public health. We are distinctive within the NIH because of our substantial investment in sustainable biomedical information systems that make scientific literature, genomic, clinical, and other types of biomedical data readily available to those who need it. We are unique because we stimulate and support innovative research in data science and information management that transcends specific disease areas and data types.

### **Key NLM Work and Innovation**

As NLM works to achieve key objectives of our Strategic Plan — to accelerate data powered discovery and health, reach new users in new ways, and prepare a workforce for a future of data-driven research and health — we also support NIH-wide efforts to answer the call to respond to national priorities, close the gap in health disparities, and capitalize on fundamental investments. We do this through effective preservation of our valued scientific and data resources, judicious investments in extramural and intramural research, informed stewardship of Federal resources, and innovative partnerships to align priorities and leverage investments across HHS, the Federal government, and the biomedical research community.

### **Answering the call to respond to national priorities**

NLM was quick to act on multiple fronts to improve the understanding of SARS-CoV-2 and aid in the global response to the COVID-19 pandemic. In January 2020, NLM released the first fully annotated SARS-CoV-2 gene sequence to the public through our GenBank® database, the world's largest database of publicly available genetic sequences. It is a significant accomplishment to release a fully annotated digital genetic sequence of a virus within a month of its detection in a human population. Because NLM maintains extensive data repositories of nucleic acid sequences, the building blocks of genes, researchers were able to search NLM's entire Sequence Read Archive (SRA) to better understand and characterize the biological



properties of SARS-CoV-2 in record time.<sup>1</sup> NLM created a dedicated website, the SARS-CoV-2 Data Resources page,<sup>2</sup> for researchers to search, retrieve, and analyze data for more than 70,000 digital genomic sequences of the virus. Providing free access to SARS-CoV-2 viral genome sequence data assures rapid dissemination and maximum impact of these data.

We facilitate scientific discovery by connecting our data and information

resources, including linking the results of individual clinical trials to associated journal publications that summarize the study. Scientists and clinicians at NIH and around the world use our resources to accelerate the development of new approaches for COVID-19 testing, evaluate promising vaccine candidates, and demonstrate the effectiveness of new therapeutics. NLM rapidly expanded access to full-text coronavirus-related scientific journal articles through our PubMed Central® (PMC) digital archive. Following a call to action issued by the White House Office of Science and Technology Policy and science policy leaders of other nations, NLM began working with the global publishing community to make tens of thousands of research publications related to coronaviruses, and associated data, freely and immediately available to the public in forms that support automated text mining. By collaborating with publishers, scholarly societies, and leading information technology (IT) companies, we helped make a massive collection of coronavirus-related information immediately accessible to artificial intelligence (AI) and machine learning researchers to accelerate discoveries about COVID-19. To ensure rapid dissemination of comprehensive information about COVID-19 research, NLM adapted and streamlined submission procedures associated with our ClinicalTrials.gov database to prioritize the processing of coronavirus-related information, and to make study results more readily discoverable. We implemented new features to ensure that information about COVID-19 studies registered on the World Health Organization's International Clinical Trial Registry Platform are discoverable through ClinicalTrials.gov. Our intramural and extramural research programs also pivoted to address the COVID-19 public health challenge. For example, NLM intramural researchers and scientists supported by NLM grants began collaborating on wastewater surveillance for early detection of SARS-CoV-2.



From January through September 2020, scientists have submitted data for more than 70,000 SARS-CoV-2 sequences to NLM's growing public archive.

### **Closing the gap in health disparities**

NLM offers unique information resources to close the gap in health disparities. We support projects that provide health information to populations experiencing health disparities, including women with disabilities, low-income Hispanic children, and sexual and gender minorities. Through more than 8,000 organizational members of our Network of the National Library of Medicine (NNLM), we deploy community-relevant strategies to make NLM's resources known to everyone, from researchers to clinicians, to the general public. With more than 60 percent of NNLM members serving underserved communities, this network provides a trusted local platform for outreach and engagement in support of major NIH initiatives such as the *All of Us*

<sup>1</sup> [www.biorxiv.org/content/10.1101/2020.08.07.241729v1](http://www.biorxiv.org/content/10.1101/2020.08.07.241729v1)

<sup>2</sup> <https://www.ncbi.nlm.nih.gov/sars-cov-2/>

Research Program, the Helping to End Addiction Long-term<sup>SM</sup> (HEAL) initiative, and the Rapid Acceleration of Diagnostics (RADx<sup>SM</sup>) initiative addressing the COVID-19 pandemic.

NLM tools such as the NIH Common Data Elements (CDE) Repository provide researchers with access to systematic ways to characterize the age, gender, and racial and ethnic aspects of participants in research studies. Use of formal approaches to define these important demographic characteristics allows researchers to better integrate the findings of one study with those from another, and better demonstrate the role of Federally funded research in illuminating the health challenges and health outcomes facing our country. Having a systematic way to express chronological age of participants also allows NLM's ClinicalTrials.gov repository to meet NIH's public commitment and public accountability for age-inclusivity in clinical trials.

We strive to improve the diversity of the biomedical information science workforce, both within NLM and the larger scientific community. Working with faculty in our 16 university-based research training programs, NLM is improving the pipeline of young scholars from groups underrepresented in the field of biomedical information science. We are working with NLM-supported investigators to expand the engagement of scientific partners from all sectors of society. For example, one of our university-based training grants is sponsoring summer research training for undergraduates at tribal colleges based on National Academy of Sciences recommendations on course-based undergraduate research experiences. NLM supports research grants to apply computational and information approaches to reduce bias in health data and health communication. For example, one project is building and evaluating a social signal processing model to improve patient-clinician communication among low income, racially diverse patients in primary care. This research is expected to yield a new approach for the next generation of health care providers and educators; empower patients experiencing health disparities; and promote health care access, quality, and equity.

### **Capitalizing on fundamental investments and beyond**

Establishing common vocabularies across clinical care settings improves the coordination of health care and the use of clinical data for research. For more than 20 years, NLM has served as the central coordinating body for clinical terminology standards nationally. Our efforts to develop, maintain, and disseminate terminology standards for medical concepts, laboratory tests, and drug names support communication among health information systems, and provide a foundation for enhanced interoperability among electronic health records (EHRs) and other critical health care information systems. Our long-standing terminology efforts contributed to the COVID-19 response by allowing access to near-real time clinical information to guide the diagnosis, treatment, and prevention of this disease.

NLM serves as a trusted source for the world's biomedical knowledge. We adapt to changing technologies that support biomedical discovery and enhance individual and public health. Over the past year, we increased use of commercial cloud computing by moving several of our essential public-facing resources to commercial cloud platforms. We worked across NIH to ensure that commercial cloud providers have the necessary security protocols and data integrity operations to protect NLM's valued biomedical resources. Effective use of cloud computing improves efficiency, reliability, and security; it also opens new opportunities for discovery. Moving large biomedical databases such as SRA to the cloud enables explorations not possible

in smaller, on-premise computer systems. It also democratizes access for scientists, industry, and the general public to use these resources for innovation.

### **Future Initiatives**

NLM will continue to lead the development of analytics that uncover new patterns and biomedical phenomena from large genetic and literature databases, create innovative ways to reach scientists and society with trustable health information, and develop health data literacy among scientists, clinicians, librarians, and consumers. Priorities for FY 2022 include efforts to:

- **Accelerate biomedical and health data science.** Through increased investments in intramural and extramural research, NLM will create strategies that support efficient and accurate exploration of large biomedical databases and generate new analytical methods and models to gain new insights from clinical data. We will also streamline our university-based research training programs to meet the information and training needs of future biomedical informatics scholars.
- **Support public accountability and open science.** We will leverage our expertise in creating high-quality, sustainable, and secure databases that make biomedical research information and data publicly accessible, such as ClinicalTrials.gov, PMC, and SRA. We will make them accessible to support the emerging needs of scientists and the public.
- **Modernize NLM's infrastructure.** With the support and collaboration of other components of NIH, we will build a 21st century digital library that uses our collections to offer literature, data, analytical models, and new approaches to scientific communications that are accessible, sustainable, and available 24 hours a day, 7 days a week.
- **Contribute to NIH- and government-wide priorities.** We will share our scientific, policy, and program expertise in data science, data management, infrastructure, security, and workforce development to support priorities empowered by data science, AI, and open science. We will provide critical support to NIH efforts such as health across the life span and the science of aging, including understanding why and how cells age (cellular senescence).

Overall Budget Policy: The FY 2022 President's Budget request is \$474.9 million, an increase of \$12.7 million from the FY 2021 Enacted level for NLM, which was \$462.1 million. Funds are included to strengthen intramural research in computational health and computational biology and enhance NLM's most critical information services, including those that provide access to published biomedical literature, scientific and clinical data, and consumer health information. We will also provide mission critical support for NIH data science goals and data sharing policies. NLM will continue key intramural data science activities, including processing and organizing genomic data essential to understanding the novel coronavirus and updating clinical terminology standards required for interoperability of U.S. health data. NLM will award an estimated 32 new research project grants through its extramural programs. NLM will continue its outreach programs that promote access and training in the effective use of NLM resources for scientists, clinicians, patients, and the general public, while continuing to advance community engagement in the NIH *All of Us* Research Program.





The National Library of Medicine (NLM), a leader in research in biomedical informatics and data science, is the world's largest biomedical library and 1 of the 27 Institutes and Centers at NIH. NLM's research and information services support scientific discovery, health care, and public health.

### A Platform for Biomedical Discovery

The NLM strategic plan includes three goals:

- 1) Accelerate discovery and advance health through data-driven research.
- 2) Reach more people in more ways through enhanced dissemination and engagement.
- 3) Build a workforce for data-driven research and health.

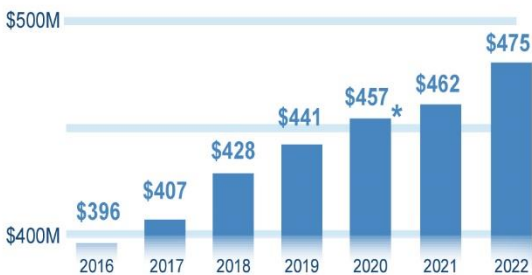
Implementation of the strategic plan enables NLM to support and respond to NIH-wide priorities.



Patricia Flatley Brennan, R.N., Ph.D.

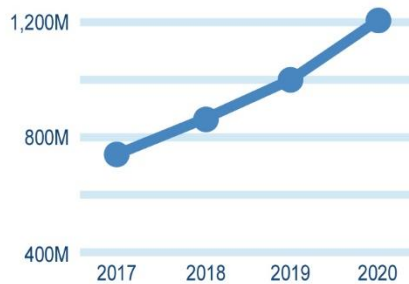
Director of the NLM since 2016, Dr. Brennan spearheaded the development of a 10-year strategic plan that envisions NLM as a platform for biomedical discovery and data-powered health. Combining her background in engineering, information technology, and clinical nursing practice, Dr. Brennan positions the NLM and its visionary biomedical data and literature resources to serve science and society, and to guide advances in data science and data-driven discovery.

### Appropriation History



\*Does not include \$10 million of supplemental funding from the Coronavirus Aid, Relief, and Economic Security (CARES) Act

### Users of NLM Services



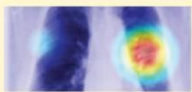
### NLM Conducts and Funds Research

NLM's vibrant intramural research program conducts innovative research and training in computational biology and computational health sciences. NLM's extramural grants support research projects and training in biomedical informatics and data science.

#### Intramural Researchers at NLM



Develop computational methods to characterize functions of the non-coding part of the human genome.



Explore novel applications of AI to image analysis, including to detect COVID-19 in chest X-rays.



Apply network analysis approaches to protein-protein interactions to predict interactions between individual molecules in human diseases.

#### Extramural Researchers Across the Nation



Leverage high-throughput sequencing data from wastewater process flows to create an innovative early warning system for virus surveillance and public health protection.



Apply machine learning to electronic health record data to support patient-matching and help clinicians explore effective treatments for similar patients.



Integrate data collected from intensive care units and create sophisticated models of causation to support health care professionals and patients in making real-time treatment decisions.





## 21st Century Library

Every day, millions of scientists, health professionals, and members of the public use NLM's online information resources to translate research results into new treatments, access our collections, develop new products, inform clinical decision making, and improve public health.



The most heavily used biomedical literature citation database in the world, containing more than 31 million citations.



Digital archive of more than 6.4 million freely accessible, full-text biomedical and life sciences journal articles.



The world's largest clinical trial registry and results database, with more than 350,000 clinical studies and 45,000 results summaries



The world's largest publicly available repository for high-throughput sequencing data, comprising more than 40 petabytes of data.



NLM's trusted and authoritative source of consumer health information, accessed by more than 371 million users annually.



NLM is the central coordinating body within HHS for clinical terminology standards for health data interoperability.

## Biomedical Informatics Training

NLM funds research training programs in biomedical informatics and data science at 16 universities across the country that enroll approximately 200 predoctoral and postdoctoral trainees.



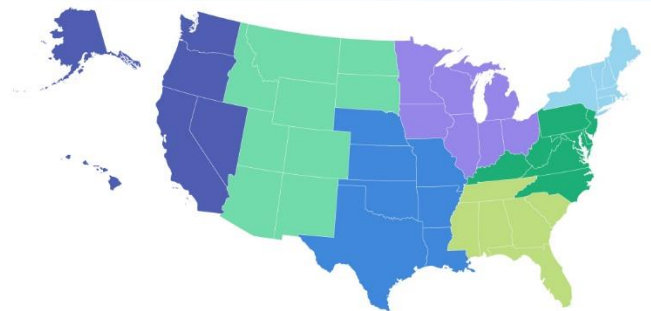
Supplemental funding allows training programs to work with library and information schools, and partner with minority-serving institutions to teach and apply data science research skills.



NLM also provides summer research experiences for undergraduates interested in careers in biomedical informatics and data science.

## Outreach and Engagement

NLM leverages its Network of the National Library of Medicine (NNLM) of more than 8,000 academic health science libraries, hospital and public libraries, and community organizations to improve access to health information for all. NNLM offers training to support effective use of NLM information resources by librarians, health professionals, researchers, and the public. NNLM will operate through seven regions across the United States from FY 2021-2026 as shown below.



## Future Initiatives

NLM will continue to lead the development of analytics that uncover novel biomedical patterns from large genetic and literature databases, create innovative ways to reach scientists and society with trustable health information, and develop health data literacy among scientists, clinicians, librarians, and consumers. NLM will:



**Accelerate biomedical and health data science.** Through investment in intramural and extramural research, NLM will create strategies that support efficient and accurate exploration of large biomedical databases and generate analytical methods and models to gain insights from clinical data.



**Modernize NLM infrastructure.** NLM will build a 21st century digital library that offers literature, data, analytical models, and new approaches to scientific communication that are accessible, sustainable, and available 24 hours a day, 7 days a week.



**Support public accountability and open science.** NLM will leverage its expertise in creating high-quality, sustainable, and secure databases and make them accessible to scientists, clinicians, and the public.



**Contribute to NIH and government-wide priorities.** NLM will share its scientific, policy, and program leadership and expertise in data science, data management, infrastructure and security, and workforce development to support the U.S. priorities in economic competitiveness empowered by data science, artificial intelligence, and open science.

## MAJOR CHANGES IN THE BUDGET REQUEST

Major changes in the FY 2022 President's Budget request for the National Library of Medicine (NLM) are briefly described below, by budget mechanism and activity detail. Note that there may be overlap between budget mechanism and activity detail; thus, these highlights will not sum to the total for NLM's FY 2022 President's Budget request, which is \$474.9 million, an increase of \$12.7 million from the FY 2021 Enacted level of \$462.1 million. Within the FY 2022 request level and informed by the *NLM Strategic Plan 2017-2027* and the *NIH Strategic Plan for Data Science*, NLM will pursue its highest priorities through strategic investments and careful stewardship of appropriated funds.

### Extramural Programs (+\$1.7 million; total \$68.0 million):

With this increased level of funding, NLM will maintain support for its 16 university-based biomedical informatics and data science training programs, which will be recompleted in FY 2022. NLM will award an estimated 32 new research project grants in biomedical informatics and data science and will maintain support the outreach and stakeholder engagement efforts of the newly configured Network of the National Library of Medicine.

### Intramural Programs (+\$10.6 million; total \$385.4 million):

NLM's intramural programs encompass intramural research and research training, as well as biomedical information services that support advances in computational health and computational biology; development of advanced biomedical information systems, standards, and research tools; acquisition, storage and distribution of biomedical data; and delivery of reliable, high-quality information services. NLM will strengthen NLM intramural research and activities that support NIH-wide interests in data science. NLM will continue to enhance its information services while providing additional support for mission critical systems that are heavily used by researchers, clinicians, and the general public. NLM will continue to support work to update and expand research use of clinical vocabularies and data interoperability standards that are increasingly important to NIH's data science efforts.

**NATIONAL INSTITUTES OF HEALTH  
National Library of Medicine**

**Budget Mechanism - Total<sup>1</sup>**

(Dollars in Thousands)

MECHANISM	FY 2020 Final		FY 2021 Enacted		FY 2022 President's Budget		FY 2022 +/- FY 2021 Enacted	
	No.	Amount	No.	Amount	No.	Amount	No.	Amount
<u>Research Projects:</u>								
Noncompeting	78	\$29,717	82	\$26,949	88	\$30,355	6	\$3,406
Administrative Supplements	(9)	649	(0)	0	(0)	0	(0)	0
<u>Competing:</u>								
Renewal	1	310	2	1,049	1	518	-1	-531
New	27	9,385	32	10,412	31	10,115	-1	-297
Supplements	0	0	0	0	0	0	0	0
Subtotal, Competing	28	\$9,695	34	\$11,462	32	\$10,634	-2	-\$828
Subtotal, RPGs	106	\$40,061	116	\$38,411	120	\$40,989	4	\$2,578
SBIR/STTR	4	1,653	5	1,553	5	1,599	0	46
Research Project Grants	110	\$41,713	121	\$39,964	125	\$42,588	4	\$2,624
<u>Research Centers:</u>								
Specialized/Comprehensive	0	\$244	0	\$32	0	\$27	0	-\$5
Clinical Research	0	0	0	0	0	0	0	0
Biotechnology	0	0	0	0	0	0	0	0
Comparative Medicine	0	0	0	0	0	0	0	0
Research Centers in Minority Institutions	0	0	0	0	0	0	0	0
Research Centers	0	\$244	0	\$32	0	\$27	0	-\$5
<u>Other Research:</u>								
Research Careers	12	\$1,646	6	\$616	2	\$186	-4	-\$430
Cancer Education	0	0	0	0	0	0	0	0
Cooperative Clinical Research	0	0	0	0	0	0	0	0
Biomedical Research Support	0	0	0	0	0	0	0	0
Minority Biomedical Research Support	0	0	0	0	0	0	0	0
Other	40	24,891	41	25,316	39	24,806	-2	-510
Other Research	52	\$26,537	47	\$25,931	41	\$24,991	-6	-\$940
Total Research Grants	162	\$68,494	168	\$65,928	166	\$67,606	-2	\$1,679
<u>Ruth L Kirschstein Training Awards:</u>	<u>FTTPs</u>		<u>FTTPs</u>		<u>FTTPs</u>		<u>FTTPs</u>	
Individual Awards	9	\$367	8	\$357	7	\$330	-1	-\$27
Institutional Awards	0	0	0	0	0	0	0	0
Total Research Training	9	\$367	8	\$357	7	\$330	-1	-\$27
Research & Develop. Contracts <i>(SBIR/STTR) (non-add)</i>	0 <i>(0)</i>	\$16 <i>(16)</i>	0 <i>(0)</i>	\$17 <i>(17)</i>	0 <i>(0)</i>	\$16 <i>(16)</i>	0 <i>(0)</i>	\$0 <i>(-0)</i>
Intramural Programs	557	367,417	638	374,847	638	385,397	0	10,550
Res. Management & Support <i>SBIR Admin. (non-add)</i>	90 <i>(0)</i>	20,616 <i>(0)</i>	103 <i>(0)</i>	20,990 <i>(0)</i>	103 <i>(0)</i>	21,514 <i>(0)</i>	0 <i>(0)</i>	525 <i>(0)</i>
Construction		0		0		0		0
Buildings and Facilities		0		0		0		0
Total, NLM	647	\$456,911	741	\$462,138	741	\$474,864	0	\$12,726

<sup>1</sup> All items in italics and brackets are non-add entries.

NATIONAL INSTITUTES OF HEALTH

National Library of Medicine

For carrying out section 301 and title IV of the PHS Act with respect to health information communications, [~~\$463,787,000~~]~~\$474,864,000~~: *Provided*, That of the amounts available for improvement of information systems, \$4,000,000 shall be available until September 30, [2022]2023: *Provided further*, That in fiscal year [2021]2022, the National Library of Medicine may enter into personal services contracts for the provision of services in facilities owned, operated, or constructed under the jurisdiction of the National Institutes of Health (referred to in this title as "NIH").



**NATIONAL INSTITUTES OF HEALTH  
National Library of Medicine**

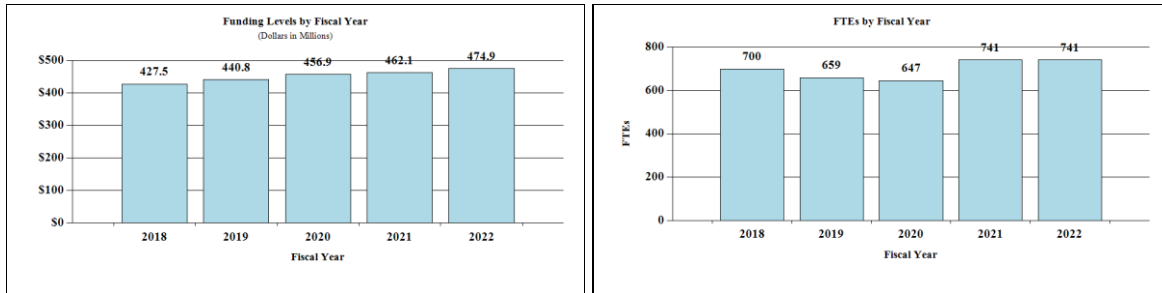
**Summary of Changes**

(Dollars in Thousands)

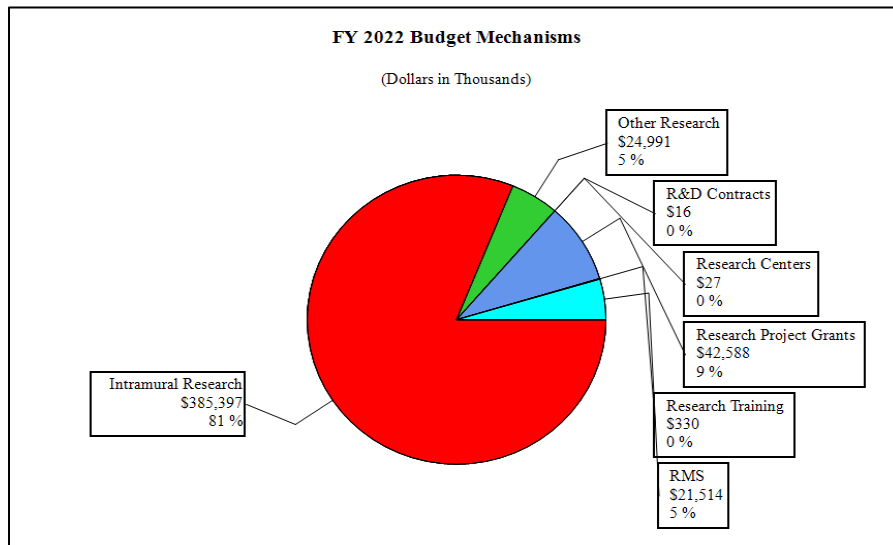
FY 2021 Enacted					\$462,138	
FY 2022 President's Budget					\$474,864	
Net change					\$12,726	
CHANGES	FY2021 Enacted		FY 2022 President's Budget		Built-In Change from FY 2021 Enacted	
	FTEs	Budget Authority	FTEs	Budget Authority	FTEs	Budget Authority
<b>A. Built-in:</b>						
<b>1. Intramural Programs:</b>						
a. Annualization of January 2021 pay increase & benefits		\$104,521		\$107,712		\$278
b. January FY 2022 pay increase & benefits		104,521		107,712		2,914
c. Paid days adjustment		104,521		107,712		0
d. Differences attributable to change in FTE		104,521		107,712		0
e. Payment for centrally furnished services		4,579		4,808		229
f. Cost of laboratory supplies, materials, other expenses, and non-recurring costs		265,748		272,878		6,124
Subtotal						\$9,545
<b>2. Research Management and Support:</b>						
a. Annualization of January 2021 pay increase & benefits		\$13,142		\$13,541		\$35
b. January FY 2022 pay increase & benefits		13,142		13,541		364
c. Paid days adjustment		13,142		13,541		0
d. Differences attributable to change in FTE		13,142		13,541		0
e. Payment for centrally furnished services		846		888		42
f. Cost of laboratory supplies, materials, other expenses, and non-recurring costs		7,002		7,086		126
Subtotal						\$567
Subtotal, Built-in						\$10,112
CHANGES	FY2021 Enacted		FY 2022 President's Budget		Program Change from FY 2021 Enacted	
	No.	Amount	No.	Amount	No.	Amount
<b>B. Program:</b>						
<b>1. Research Project Grants:</b>						
a. Noncompeting	82	\$26,949	88	\$30,355	6	\$3,406
b. Competing	34	11,462	32	10,634	-2	-828
c. SBIR/STTR	5	1,553	5	1,599	0	46
Subtotal, RPGs	121	\$39,964	125	\$42,588	4	\$2,624
2. Research Centers	0	\$32	0	\$27	0	-\$5
3. Other Research	47	25,931	41	24,991	-6	-940
4. Research Training	8	357	7	330	-1	-27
5. Research and development contracts	0	17	0	16	0	0
Subtotal, Extramural		\$66,301		\$67,953		\$1,652
6. Intramural Programs	<u>FTEs</u> 638	\$374,847	<u>FTEs</u> 638	\$385,397	<u>FTEs</u> 0	\$1,005
7. Research Management and Support	103	20,990	103	21,514	0	-42
8. Construction		0		0		0
9. Buildings and Facilities		0		0		0
Subtotal, Program	741	\$462,138	741	\$474,864	0	\$2,614
Total built-in and program changes						\$12,726

## Fiscal Year 2022 Budget Graphs

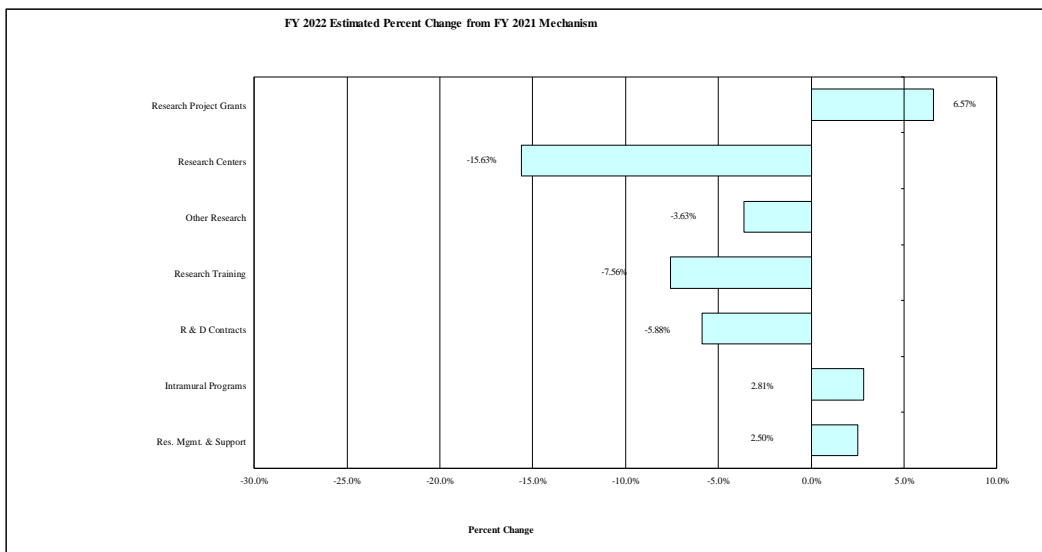
History of Budget Authority and FTEs:



Distribution by Mechanism:



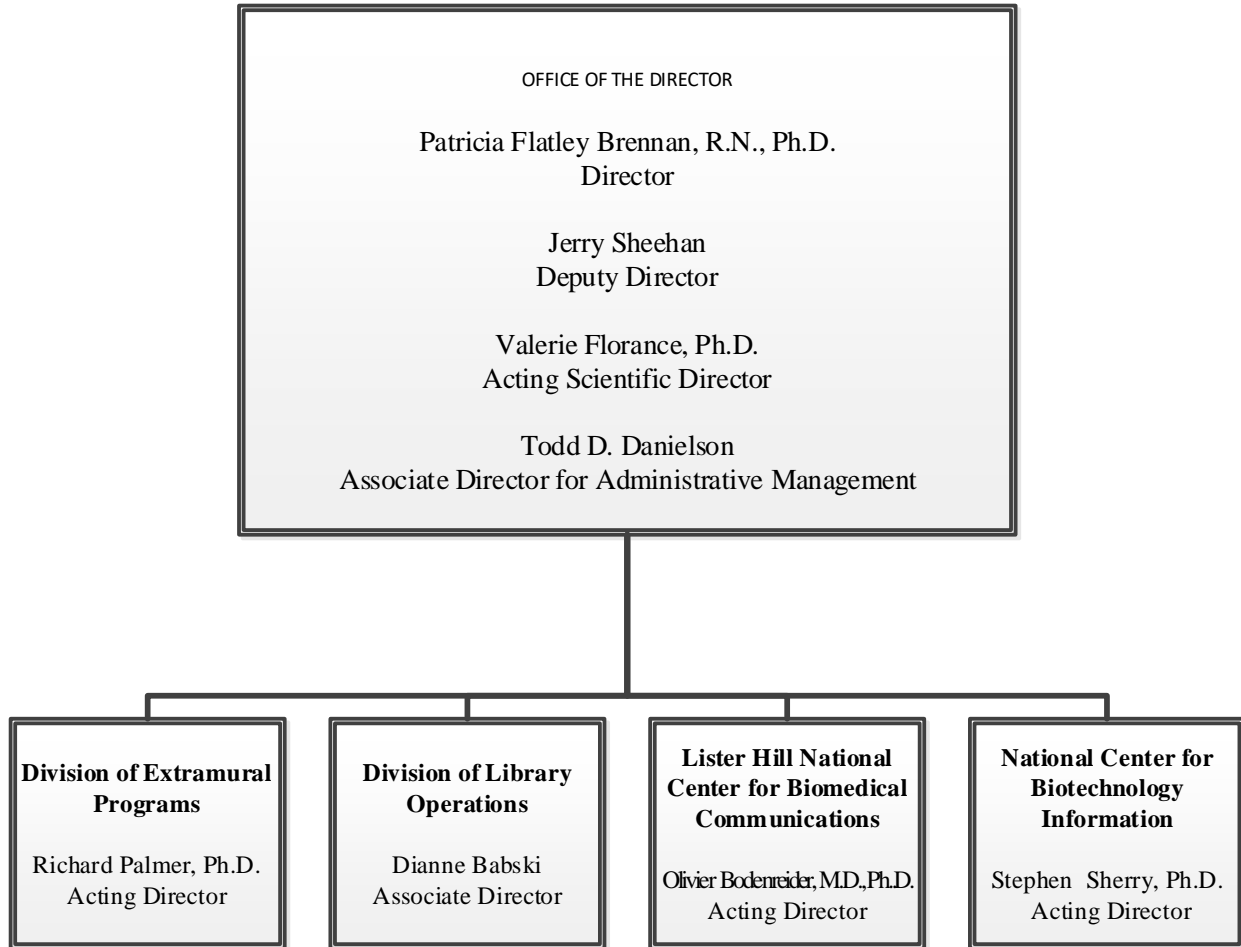
Change by Selected Mechanism:



NATIONAL INSTITUTES OF HEALTH

National Library of Medicine

ORGANIZATIONAL CHART



**NATIONAL INSTITUTES OF HEALTH**  
**National Library of Medicine**

**Budget Authority by Activity<sup>1</sup>**  
(Dollars in Thousands)

	FY 2020 Final		FY 2021 Enacted		FY 2022 President's Budget		FY 2022 +/- FY 2021 Enacted	
	FTE	Amount	FTE	Amount	FTE	Amount	FTE	Amount
<b>Extramural Research</b>								
<u>Detail</u>								
Health Information for Health Professionals and the Public		\$11,132		\$11,559		\$11,559		\$0
Informatics Resources for Biomedicine and Health		17,469		14,778		13,806		-972
Biomedical Informatics Research		40,276		39,964		42,588		2,624
<b>Subtotal, Extramural</b>		<b>\$68,878</b>		<b>\$66,301</b>		<b>\$67,953</b>		<b>\$1,652</b>
<b>Intramural Programs</b>	<b>557</b>	<b>\$367,417</b>	<b>638</b>	<b>\$374,847</b>	<b>638</b>	<b>\$385,397</b>	<b>0</b>	<b>\$10,550</b>
<b>Research Management &amp; Support</b>	<b>90</b>	<b>\$20,616</b>	<b>103</b>	<b>\$20,990</b>	<b>103</b>	<b>\$21,514</b>	<b>0</b>	<b>\$525</b>
<b>TOTAL</b>	<b>647</b>	<b>\$456,911</b>	<b>741</b>	<b>\$462,138</b>	<b>741</b>	<b>\$474,864</b>	<b>0</b>	<b>\$12,726</b>

<sup>1</sup> Includes FTEs whose payroll obligations are supported by the NIH Common Fund.

## Justification of Budget Request

### National Library of Medicine

Authorizing Legislation: Section 301 and title IV of the Public Health Service Act, as amended.

Budget Authority (BA):

	FY 2020 Final	FY 2021 Enacted	FY 2022 President's Budget	FY 2022 +/- FY 2021
BA	<u>\$456,911,000</u>	\$462,138,000	\$474,864,000	+\$12,726,000
FTE	647	741	741	0

Program funds are allocated as follows: Competitive Grants/Cooperative Agreements; Contracts; Direct Federal/Intramural and Other.

### Program Descriptions and Accomplishments

#### Intramural Programs

NLM's intramural programs encompass two major activities: 1) Intramural Research and Training; and 2) Biomedical Information Services.

#### **Intramural Research and Training**

NLM's intramural research program is distinctive because it focuses on solving biological and clinical problems through the development of computational approaches that can be generalized and applied across multiple types of data, disciplines, and health issues. NLM conducts research and training in computational biology and computational health sciences, leverages its robust biomedical information resources to answer complex questions with novel analytical approaches, and validates those methods for general use. This work generates new knowledge, methods, and tools to help scientists and clinicians better understand basic biology and the natural history of disease, discover new approaches to care, and improve clinical decisions and health outcomes. For example, an NLM researcher developed a new computational method to classify viruses and then, in collaboration with the International Committee on the Taxonomy of Viruses, disseminated the resultant taxonomy through NLM production resources where it is accessible to researchers around the world.

NLM's computational biology research develops and applies advanced algorithms and models to explore molecular and image data to better understand biological systems, and to examine biological phenomena across many levels of organization: chemical, molecular, cellular, organ, individual, or population. Recent work has explored the physical characteristics of proteins, examining the ways these basic building blocks of life change their structures and functions in response to changes in their environment. Other work examining basic biological structures has resulted in a new computational approach to map what was once thought of as a non-functional part of the human genome and has since been determined to regulate how other genes are turned



on and off. These types of biological discoveries create new opportunities to determine where gene modifications or therapies may be effective. Other research employs a multiple-organism approach to devise an innovative strategy based on comparative genomics to examine clusters of single-cell microbes to identify how microbes in close proximity share features with each other. Once these similarities are identified computationally, investigators can generate hypotheses that can be tested and confirmed in biological laboratories. NLM researchers also use network analysis to detect tumor clusters by identifying associations between gene structure mutations in cells, which could help researchers and clinicians understand more about how tumors evolve and improve methods of treatment. This innovative work was codified in software that assists in the design and development of novel drug design pipelines.

NLM's computational health research focuses on clinical information processing. It includes research on computational methods for mining clinical data, analyzing human images, and automating the development of new health data standards. Researchers have used analytical tools powered by those health data standards to explore health claims databases and reveal new patterns of drug responses, such as the unanticipated benefit of estrogen replacement therapy in women who have had hysterectomies. Investigators have used natural language processing to better understand and respond to the needs of stakeholders who use NLM's literature and consumer health information resources. NLM researchers have developed AI and machine learning approaches to examine hundreds of clinical images to automate screening for diseases such as cervical cancer and advanced macular degeneration.

In FY 2020, NLM's intramural researchers quickly realigned their efforts to respond to the COVID-19 pandemic by adapting and enhancing image analysis, terminology research, natural language processing tools, and AI techniques to:

- Identify lung abnormalities in chest X-rays to distinguish between pneumonia caused by a bacteria or virus, and identify unique visual features associated with COVID-19.
- Understand the biology and development of the SARS-CoV-2 virus and related coronaviruses and identify fast-evolving regions of SARS-CoV-2 proteins that may affect how the immune system responds to this virus.

### **Advancing Research in Artificial Intelligence**

NLM plays an enormous role in supporting AI research and its application to biomedical discovery and clinical care at NIH. NLM researchers explore several domains of AI research, including:

- Natural language processing to develop and evaluate algorithms for automated analysis of text such as biomedical literature or electronic health records (EHRs). Recent work contributed to development of the Best Match search algorithm used in NLM's PubMed biomedical literature database, and enabled rapid curation of LitCovid, a literature hub with up-to-date scientific publications related to the novel coronavirus.
- Image processing to support clinical decision-making. NLM researchers developed and applied novel AI approaches to the screening, diagnosis, and prognosis of many diseases and disorders including age-related macular degeneration, cervical cancer, and drug-resistant tuberculosis.
- Machine learning to conduct DNA sequence analysis. NLM researchers used AI to identify 2,500 families of proteins that can potentially be used to regulate the activity of gene editing tools based on CRISPR-Cas, a technique that holds tremendous promise for gene therapy and scientific research.

In FY 2020, NLM's extramural program supported 11 AI-related research projects to help curate large-scale datasets and predictive analytics from EHRs, clinical decision-making for organ donation, and autism therapy for children. NLM also led a collaboration with the National Institute of Standards and Technology to launch team challenges, including the COVID Epidemic Question Answering.

- Facilitate searching for coronavirus-related images and articles in biomedical literature.
- Mine claims and hospital data associated with COVID-19 to gain new insight on the disease.

NLM's intramural research program offers a vibrant training environment that includes laboratory-based mentored training, short-term rotations for clinicians, and advanced weekly seminars and lectures. NLM offers training for scholars in biomedical informatics, computational biology, and data science. In FY 2020, NLM hired 2 new tenure-track investigators with expertise in computational approaches to characterize structures of proteins and viruses, and hosted 46 postdoctoral researchers, 5 postbaccalaureate trainees, and 2 visiting scientists. NLM also launched the Ada Lovelace Computational Health Lecture Series to provide a forum to explore and increase attention to the contributions of computational innovation to biomedical research and health.

### **Biomedical Information Services**

NLM's biomedical information services support scientific discovery and innovation through the collection, dissemination, and exchange of data and information critical to medicine and health. Each day, more than 7 million people and computer information systems use NLM's web-based biomedical information services and download 270 terabytes of data. NLM accepts and processes more than 13 terabytes of data submitted by thousands of researchers and businesses every day. Continually rising demand necessitates that NLM regularly upgrade its biomedical information services to improve operational efficiencies and accommodate the growing volume and variety of data and information needed to support research and discovery. In FY 2020, NLM enhancements included efforts to migrate services to modern computing platforms and improve data security. NLM also made notable progress in advancing its work in health data standards, as well as outreach and engagement to promote use of its resources.

### Improving Access to Biomedical Literature.

NLM is a primary resource for millions of researchers, clinicians, students, and members of the general public to find and access biomedical literature. Users can access NLM's flagship database of citations to the biomedical literature via mobile or desktop devices using NLM's re-engineered PubMed®, which includes an improved search engine driven by AI, responsive design to optimize features for mobile devices, and an improved user interface that increases control over the ways the results are displayed. An important innovation is the display of results ranked by relevance, rather than by publication date. NLM also added 1.4 million citations to PubMed in FY 2020, increasing the total collection to more than 31 million bibliographic citations.

NLM continues to serve as a leader in ensuring free public access to the results of biomedical research through PubMed Central (PMC), its full-text archive of biomedical literature. NLM added 700,000 full-text articles to PMC in FY 2020 and continues to enhance the value of these articles by linking them to associated datasets. Of the 6.4 million articles in PMC at the end of FY 2020, more than 1.5 million included associated data, and more than 3.5 million were downloadable in machine-readable forms to support automated analysis. Ten other Federal agencies continue to use PMC as the most efficient and effective means to provide free, public access to full-text journal articles reporting on research they fund. To respond to the COVID-19

pandemic, NLM established new partnerships with publishers to expand access to journal articles reporting on coronavirus-related research. NLM leveraged this archive to launch a pilot program to make prepublication (preprint) versions of articles available and searchable through PMC, focusing on preprints resulting from NIH-funded research on COVID-19.

NLM also maintains its Digital Collections, a free, online repository of historical and modern biomedical resources including books, manuscripts, still images, videos, sound recordings, and maps that serve more than 180,000 librarians, researchers, and others each year. In FY 2020, NLM deployed innovative techniques to prospectively curate and add COVID-19-related information from traditional news, social media, and other sources to the Digital Collections.

#### Molecular Biology Data and Tools

NLM maintains an array of more than 40 integrated molecular biology databases and bioinformatics tools that are freely accessible and enable biomedical research and discovery. Notably, in FY 2020 NLM completed the transition of more than 40 petabytes of genetic sequence data in multiple formats from the SRA repository to secure commercial cloud providers participating in the NIH Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability (STRIDES) Initiative. Not only is cloud-based storage cost-effective, it provides access to and preservation of genomic sequence data in a secure, sustainable manner. Storing genomic data safely in the cloud also has the potential to transform research by offering scientists the ability to compute across the complete corpus of high-throughput sequence data without having to download it — something not possible using traditional data management approaches. One research team has already explored the corpus of genome sequence data for signatures of coronavirus. This analysis, which previously would have taken more than a year to complete, was accomplished in three days and delivered tens of thousands of coronavirus gene alignments freely to the research community to catalyze a new era of virus discovery.

#### **Information Access in a Pandemic**

NLM responds to scientific challenges by providing access to biomedical literature and data. Since the pandemic began, NLM developed a centralized SARS-CoV-2 webpage to facilitate access to submission interfaces and content from NLM resources such as GenBank and SRA (sequence data), PubMed and PMC (literature), PubChem® (chemical data), and ClinicalTrials.gov. In March 2020, NLM launched its Public Health Emergency COVID-19 Initiative, bringing together publishers and scientific scholarly societies to make COVID-19 and coronavirus-related articles freely accessible through PMC in machine-readable formats that permit research re-use and analysis.

By the end of FY 2020, NLM made available more than 80,000 articles which were accessed more than 40 million times by people and computer systems. NLM contributed this collection to the larger COVID-19 Open Research Dataset (CORD-19), developed in collaboration with private-sector partners and used in AI-based challenges to address key scientific questions related to COVID-19. NLM also facilitated early access to research results from NIH-funded research with the launch of the NIH Preprint Pilot in June 2020. This pilot tests the viability of making article preprints searchable in PMC and discoverable in PubMed, beginning with COVID-19 preprints. By the end of FY 2020, nearly 1,000 preprints were added to these databases; 80 percent were still unavailable in published journals.

To coordinate U.S. efforts to provide publicly accessible SARS-CoV-2 sequence data, NLM contributes to the SARS-CoV-2 Sequencing for Public Health Emergency Response, Epidemiology and Surveillance (SPHERES) consortium. NLM improved access to gene sequence information related to SARS-CoV-2, making the first sequence available in GenBank in January 2020, and established the COVID-19 Genome Sequence Dataset, providing free cloud-based access to SARS-CoV-2 SRA data through the NIH cloud-based STRIDES Initiative. NLM developed a specialized SARS-CoV-2 virus resource for searching and analyzing sequences; streamlined data submission processes; and integrated data validation to support research and accelerate the pandemic response.

NLM also created the *COVID-19 Genome Sequence Dataset* within the SRA to provide scientists with cloud-based access to SARS-CoV-2 SRA data.

NLM resources must keep pace with advances in science and expanding knowledge of genomics. NLM continually improves its molecular biology resources to ensure comprehensive data collection, including content related to coronaviruses. For example, NLM added data for 470 million genetic sequences to GenBank, which describes all publicly available DNA sequences, including sequences for SARS-CoV-2. NLM added 40 million records (a 19 percent increase) to the Reference Sequence (RefSeq) collection, which provides a comprehensive collection of reference sequences and genes against which individual variations can be compared. NLM added 100,000 human genome sequence variants to ClinVar, which provides researchers with aggregated information about clinically significant genomic variations among humans. NLM facilitated submission of 210 studies to the database of Genotypes and Phenotypes (dbGaP), which archives and provides controlled access to data from large-scale studies of the relationship between genetic characteristics and disease, and physical characteristics (phenotypes). New analyses of study data made available by dbGaP resulted in more than 3,000 research publications by the end of FY 2020. In collaboration with public health authorities and the Food and Drug Administration (FDA), NLM processed genome sequence data for 197,000 samples through its Pathogen Detection Pipeline to help identify sources of illness resulting from foodborne pathogens such as Salmonella, E. coli, and Listeria. Using the pipeline, NLM has supported the FDA in more than 660 actions to protect consumers from foodborne illness.

#### Clinical Trials Information

The COVID-19 pandemic underscores the importance of having a gold standard of research evidence for clinical care. NLM's ClinicalTrials.gov is the world's largest publicly accessible database of privately and publicly funded clinical studies. In FY 2020, NLM added information about more than 35,000 new clinical research studies and 6,500 new results summaries. Making these visible to the scientific community and the public improves public accountability and increases public trust in science. To support the global response to COVID-19 and ensure streamlined access to up-to-date information about related studies conducted around the world, NLM added more than 4,000 clinical studies, including 2,800 listed with the World Health Organization. NLM implemented new procedures to ensure that submitted results information is posted no later than 30 days after submission. NLM also launched a major ClinicalTrials.gov modernization effort and engaged more than 300 individuals and organizations to provide public input about expected and desired features for the modernized system.

#### Consumer Health Information

While NLM's biomedical information services are accessible to the public, NLM's MedlinePlus® is specifically designed to meet consumers' health information needs. Available in English and Spanish, MedlinePlus offers trusted, authoritative information for patients and families about a broad variety of health conditions, medical tests, drugs and supplements, recipes, and videos, along with links to other credible sources of information. More than 371 million users accessed MedlinePlus in FY 2020. To provide a more comprehensive, consolidated consumer-oriented resource, NLM integrated its popular Genetics Home Reference™ website into MedlinePlus to provide information about genetic conditions, genes, and genetic variations. NLM's MedlinePlus Connect service enables EHRs, patient portals, and

other healthcare IT systems to deliver information from MedlinePlus to patients and providers at the point of care. In FY 2020, MedlinePlus Connect seamlessly responded to more than 217 million electronic requests from health IT systems.

### Standards and Terminologies for Health Data Interoperability

NLM works across the NIH, the Federal government, and the health care delivery system to support the development, maintenance, and dissemination of health data standards that promote interoperability among EHRs, as well as other clinical and research information systems. In FY 2020, NLM supported expansion of three clinical language systems used for clinical terminology, laboratory tests and observations, and drug names to include terms and codes associated with SARS-CoV-2 and COVID-19. NLM released an updated U.S. edition of the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT<sup>®</sup>) with 3,800 new concepts, 40 of which are COVID-19-related. NLM added more than 1,200 new terms to the Logical Observation Identifiers, Names and Codes (LOINC<sup>®</sup>) for laboratory tests and observations, 160 of which are COVID-19-related. NLM updated drug names in RxNorm to include select investigational drugs related to COVID-19. Additional enhancements in FY 2020 included automated procedures to make it easier for clinicians and researchers to measure quality of care related to COVID-19. NLM continues to provide and enhance NIH's CDE Repository, a freely available source of standard, structured, machine-readable definitions of data elements, standard variables, and measures used in NIH-funded clinical research. In FY 2020, NLM updated and added new CDEs and forms to the repository from six NIH Institutes, Centers, and initiatives. NLM introduced enhancements to the CDE Repository to improve the user experience and facilitate COVID-19 data services.

### Outreach and Engagement

NLM's outreach and engagement activities raise awareness and support effective use of its information, data resources, and physical collections. Much of this work is conducted through the NNLM, which leverages more than 8,000 academic health science libraries, hospital and public libraries, and community organizations to reach people across the nation, including in medically underserved communities. Collaborations between communities and researchers build capacity to address problems and meet research goals. Community participation in the research process also builds trust between NNLM and the communities that we serve. NNLM programs promote digital and health literacy, support the NIH *All of Us* Research Program and its efforts to engage diverse

#### **Health Data Standards for Research**

NLM is a pioneer in clinical terminology and messaging standards to support nationwide interoperability of health information systems. In FY 2020, NLM made notable headway in expanding the use of health data standards in clinical research. NLM supports the development, use, and usability of a data exchange standard to improve the availability of clinical data for research via the Health Level Seven International (HL7) Fast Healthcare Interoperability Resources (FHIR) standard. NLM manages the development and testing of FHIR tools that researchers can use to: increase the availability of high-quality, standardized research datasets and phenotypic information for genomic research and genomic medicine; capture, integrate, and exchange clinical data for research purposes; and enhance capabilities to share research data. NLM collaborates across NIH and the Federal government to understand what additional development is needed to effectively leverage the FHIR standard in research.

NLM leads NIH efforts to encourage NIH-supported researchers to adopt and use content and vocabulary standards specified in the United States Core Data for Interoperability, promulgated by the HHS Office of the National Coordinator for Health Information Technology. This set of standards can facilitate the use of clinical data in research studies and foster greater consistency across clinical research data shared with others.



communities, enhance the skills of health sciences librarians in research data management and data science, and encourage community science. In FY 2020, 4,000 NNLM citizen science and crowdsourcing activities engaged 1.5 million people in research to address societal needs and accelerate biomedical science, technology, and innovation. NLM also supports engagement through exhibitions and seminars that highlight NLM collections and resources. In FY 2020, NLM and the National Endowment for the Humanities cosponsored a virtual research symposium on aspects of the 1918 influenza pandemic that could inform responses to the COVID-19 pandemic.

**Budget Policy:** The FY 2022 President's Budget estimate for NLM's Intramural Programs is \$385.4 million, an increase of \$10.6 million from the FY 2021 Enacted Budget of \$374.8 million. NLM will continue to revitalize its intramural research program as it improves coordination of research activities under a single Scientific Director and seeks efficiencies in core research services. Research priorities will be informed by the needs of NLM, NIH, and the broader biomedical research community, including for novel artificial intelligence and machine learning methods for analyzing large clinical and biological data sets; improved approaches for developing and applying health data standards; and automated techniques for processing, organizing, indexing, and managing biomedical literature and data. NLM will prioritize and seek additional efficiencies in its information services, including through use of cloud-based services, to ensure needed levels of support for those that are most heavily used by scientific researchers, clinicians, and the public and that support NIH-wide data science goals and data sharing policies. NLM will continue its modernization of ClinicalTrials.gov to facilitate submission of and access to clinical trial data submitted in accordance with the Food and Drug Administration Amendments Act of 2007 and NIH policy. It will improve tools for updating and disseminating clinical terminology standards that support interoperability of electronic health data and advanced integration and analysis of genomic, clinical research, and observational health data. NLM will prioritize its outreach programs that promote access and training in the effective use of NLM resources, including data repositories, and to increase engagement with broad sets of stakeholders in the public and private sectors to leverage their talents, capabilities, and information resources.

### **Extramural Programs**

NLM's extramural programs encompass three major activities: 1) Biomedical Informatics Research; 2) Informatics Resources for Biomedicine and Health; and 3) Health Information for Health Professionals and the Public.

### **Biomedical Informatics Research**

NLM's research grants address the growing demand for innovation in computational and data science approaches to make data findable, accessible, interoperable, and reusable. In FY 2020, NLM funded 186 awards to investigators across the country, including 6 co-funded with other NIH Institutes and Centers. NLM issued Notices of Special Interest seeking applications in two major areas: 1) mining clinical data to understand how a disease presents in an individual to identify or predict the presence of COVID-19; and 2) public health surveillance methods that mine a variety of data types to identify the spread and impact of SARS-CoV-2. Additional funding in FY 2020, \$10.0 million from the Coronavirus Aid, Relief, and Economic Security

(CARES) Act, enabled NLM to provide supplemental funds to support new methods to rapidly improve the understanding of SARS-CoV-2 and COVID-19.

NLM-funded researchers develop methods to capture, analyze, visualize, integrate, and curate biomedical data to support discovery and decision-making. Active NLM-supported research projects include developing statistical methods for clinical data analysis and applying AI to pulmonary imaging, computer-aided diagnosis for brain cancer, and new machine learning methods. In FY 2020, NLM-funded researchers:

- Used data from EHRs and clinical databases to predict: the glycemic impact of meals to identify patients with glucose intolerance, disease trajectory of hospitalized COVID-19 patients compared to past hospitalized flu patients, and risk for pulmonary embolism by using computed tomography imaging and patient-specific risk factors.
- Used natural language processing to characterize trajectories of symptoms, interventions, and treatment response from free text in the EHR, and create patient timelines for more precise estimates of potential medication interactions.
- Generated and evaluated new computer-aided diagnosis techniques to assist in the clinical detection of cancer metastases in the brain using magnetic resonance imaging, which has the potential to increase the accuracy of early detection, allow prompt treatment, and improve patients' well-being.
- Developed AI methods to integrate data from intensive care units to guide real-time treatment decisions and support shared decision-making using a new statistical algorithm that integrates patients' values, preferences, and post-treatment outcomes.

#### University-based training program for biomedical informatics and data science

NLM is a leading funder of Ph.D.-level training in biomedical informatics and data science. NLM's university-based research training program supports 16 universities across the country that enroll approximately 200 predoctoral and postdoctoral fellows and trainees. The programs leverage NLM's long-term investment strategy to help shape the field of biomedical informatics and data science. Graduates serve as leaders in academia and industry. NLM encourages its university-based programs to share curriculum materials and use-case examples with other graduate-level training programs to accelerate data science training in other universities. NLM provides awards to sponsor summer research experiences for undergraduates and others interested in careers in biomedical informatics and data science. In FY 2020, supplemental funding enabled training programs to work with a coalition of 12 faculty from minority-serving institutions who designed a 10-week data science research internship program incorporating data science skills focused on machine learning, statistics, and clinical informatics; ethical considerations in data science; and core skills for academic success and career development related to the analysis and presentation of scientific literature and findings. In FY 2020, NLM supported three NIH National Research Service Award predoctoral fellowships.

#### **Informatics Resources for Biomedicine and Health**

NLM's Information Resource Grants to Reduce Health Disparities program supports projects that bring useful and understandable health information to populations affected by health disparities, and their health care providers. One team is developing an AI tool to provide visual health information to help communities prevent spread of antibiotic resistant infections. NLM resource grants encourage collaboration between Historically Black Colleges and Universities,

Tribal Colleges, other Minority-Serving Institutions, and biomedical informatics training sites. NLM's Grants for Scholarly Works in Biomedicine and Health support works for health professionals, public health officials, biomedical researchers, and health science historians. In FY 2020, NLM awarded two grants for scholarly works about Sibling Obligations in Health Care, and Jim Crow in the Asylum: Psychiatry and Civil Rights in the American South.

### **Health Information for Health Professionals and the Public**

The NLM Extramural Program oversees the cooperative agreements that fund NNLM to improve the health of the Nation by training U.S. health professionals to access biomedical information, enhancing individuals' access to trusted health information to make decisions about their health, and building capacity for data management and science at health sciences libraries.

Budget Policy: The FY 2022 President's Budget request includes \$68.0 million for NLM's Extramural Programs, an increase of \$1.7 million from the FY 2021 Enacted level of \$66.3 million. NLM will continue to accept investigator-initiated applications through NIH parent-grant announcements, as well as applications submitted to NLM's own funding announcements. In FY 2022, NLM will support noncompeting grants at the previously committed level. NLM expects to award an estimated 32 competing research project grants and will aim to support early stage and new investigators at success rates comparable to those of established investigators submitting new applications. It will continue to support university-based research training programs in biomedical informatics and data science at up to 16 universities across the country that enroll approximately 200 predoctoral and postdoctoral trainees. NLM will continue to support its unique resource grant programs and career transition programs. Through its cooperative agreement for the Network of the National Library of Medicine, NLM will continue to support efforts to advance community engagement in the *All of Us* Research Program, and to retain support for its engagement and training programs.

### **Research Management and Support**

NLM's research management and support (RMS) activities provide administrative, budgetary, communications, and logistical support for NLM programs to ensure strategic planning and evaluation; regulatory compliance; policy development; international coordination; and partnerships with other Federal agencies, Congress, the private sector, and the public. In FY 2020, NLM streamlined its organizational and administrative structure to enhance collaborative leadership, innovation, and customer service. NLM consolidated management of its engagement and training initiatives into a single office, and unified management of its intramural research program under a single scientific director.

Budget Policy: The FY 2022 President's Budget request includes \$21.5 million for NLM's RMS activities, an increase of \$0.5 million from the FY 2021 Enacted level of \$21.0 million. RMS will support NLM-wide planning and evaluation, including implementation and updating of NLM's strategic plan. It will also support enhancement of NLM's critical physical and information systems security infrastructure, policy development and administration functions, and improved coordination of trans-NLM and trans-NIH efforts in data science.

**Conclusion:** Through its research, information systems, and engagement activities, NLM serves as a platform for biomedical data discovery and data-powered health, enabling researchers,

clinicians, and the public to use the vast wealth of biomedical knowledge to improve the health of the Nation.

**NATIONAL INSTITUTES OF HEALTH  
National Library of Medicine**

**Appropriations History**

<b>Fiscal Year</b>	<b>Budget Estimate to Congress</b>	<b>House Allowance</b>	<b>Senate Allowance</b>	<b>Appropriation</b>
2013	\$372,651,000		\$381,981,000	\$337,638,655
Rescission				\$675,277
Sequestration				(\$16,947,139)
2014	\$382,252,000		\$387,912,000	\$327,723,000
Rescission				\$0
2015	\$372,851,000			\$336,939,000
Rescission				\$0
2016	\$394,090,000	\$341,119,000	\$402,251,000	\$394,664,000
Rescission				\$0
2017 <sup>1</sup>	\$395,684,000	\$407,086,000	\$412,097,000	\$407,510,000
Rescission				\$0
2018	\$373,258,000	\$413,848,000	\$420,898,000	\$428,553,000
Rescission				\$0
2019	\$395,493,000	\$433,671,000	\$442,230,000	\$441,997,000
Rescission				\$0
2020	\$380,463,000	\$463,599,000	\$465,837,000	\$456,911,000
Rescission				\$0
Supplemental				\$10,000,000
2021	\$415,665,000	\$460,841,000	\$471,789,000	\$463,787,000
Rescission				\$0
2022	\$474,864,000			

<sup>1</sup> Budget Estimate to Congress includes mandatory financing.

**NATIONAL INSTITUTES OF HEALTH**  
National Library of Medicine

**Authorizing Legislation**

	<b>PHS Act/ Other Citation</b>	<b>U.S. Code Citation</b>	<b>2021 Amount Authorized</b>	<b>FY 2021 Enacted</b>	<b>2022 Amount Authorized</b>	<b>FY 2022 President's Budget</b>
Research and Investigation	Section 301	42§241	Indefinite	\$462,138,000	Indefinite	\$474,864,000
National Library of Medicine	Section 401(a)	42§281	Indefinite		Indefinite	
<b>Total, Budget Authority</b>				<b>\$462,138,000</b>		<b>\$474,864,000</b>

**NATIONAL INSTITUTES OF HEALTH  
National Library of Medicine**

**Amounts Available for Obligation<sup>1</sup>**  
(Dollars in Thousands)

Source of Funding	FY 2020 Final	FY 2021 Enacted	FY 2022 President's Budget
Appropriation	\$456,911	\$463,787	\$474,864
Secretary's Transfer	0	0	0
OAR HIV/AIDS Transfers	0	-1,649	0
Subtotal, adjusted budget authority	\$456,911	\$462,138	\$474,864
Unobligated balance, start of year	2,500	2,620	0
Unobligated balance, end of year	-2,620	0	0
Subtotal, adjusted budget authority	\$456,791	\$464,758	\$474,864
Unobligated balance lapsing	-207	0	0
Total obligations	\$456,584	\$464,758	\$474,864

<sup>1</sup> Excludes the following amounts (in thousands) for reimbursable activities carried out by this account:  
FY 2020 - \$7,706    FY 2021 - \$8,000    FY 2022 - \$8,080



**NATIONAL INSTITUTES OF HEALTH**  
**National Library of Medicine**

**Budget Authority by Object Class<sup>1</sup>**  
(Dollars in Thousands)

	<b>FY 2021 Enacted</b>	<b>FY 2022 President's Budget</b>	<b>FY 2022 +/- FY 2021 Enacted</b>
Total compensable workyears:			
Full-time equivalent	741	741	0
Full-time equivalent of overtime and holiday hours	1	1	0
Average ES salary	\$199	\$204	\$5
Average GM/GS grade	11.8	11.8	0.0
Average GM/GS salary	\$111	\$114	\$3
Average salary, Commissioned Corps (42 U.S.C. 207)	\$120	\$123	\$3
Average salary of ungraded positions	\$159	\$162	\$4
<b>OBJECT CLASSES</b>	<b>FY 2021 Enacted</b>	<b>FY 2022 President's Budget</b>	<b>FY 2022 +/- FY 2021</b>
Personnel Compensation			
11.1 Full-Time Permanent	42,248	43,209	961
11.3 Other Than Full-Time Permanent	41,808	42,759	951
11.5 Other Personnel Compensation	1,458	1,491	33
11.7 Military Personnel	120	123	3
11.8 Special Personnel Services Payments	1,888	1,931	43
<b>11.9 Subtotal Personnel Compensation</b>	<b>\$87,522</b>	<b>\$89,514</b>	<b>\$1,992</b>
12.1 Civilian Personnel Benefits	30,052	31,647	1,596
12.2 Military Personnel Benefits	89	92	2
13.0 Benefits to Former Personnel	0	0	0
<b>Subtotal Pay Costs</b>	<b>\$117,663</b>	<b>\$121,253</b>	<b>\$3,590</b>
21.0 Travel & Transportation of Persons	420	428	8
22.0 Transportation of Things	308	313	6
23.1 Rental Payments to GSA	165	168	3
23.2 Rental Payments to Others	140	143	3
23.3 Communications, Utilities & Misc. Charges	477	486	9
24.0 Printing & Reproduction	81	82	1
25.1 Consulting Services	87,601	89,351	1,750
25.2 Other Services	70,790	73,563	2,773
25.3 Purchase of goods and services from government accounts	72,411	74,526	2,115
25.4 Operation & Maintenance of Facilities	213	213	0
25.5 R&D Contracts	17	16	0
25.6 Medical Care	1	1	0
25.7 Operation & Maintenance of Equipment	14,689	14,954	264
25.8 Subsistence & Support of Persons	5	5	0
<b>25.0 Subtotal Other Contractual Services</b>	<b>\$245,727</b>	<b>\$252,630</b>	<b>\$6,903</b>
26.0 Supplies & Materials	1,189	1,210	21
31.0 Equipment	29,177	29,702	525
32.0 Land and Structures	335	342	6
33.0 Investments & Loans	0	0	0
41.0 Grants, Subsidies & Contributions	66,452	68,104	1,652
42.0 Insurance Claims & Indemnities	0	0	0
43.0 Interest & Dividends	3	3	0
44.0 Refunds	0	0	0
<b>Subtotal Non-Pay Costs</b>	<b>\$344,475</b>	<b>\$353,611</b>	<b>\$9,136</b>
<b>Total Budget Authority by Object Class</b>	<b>\$462,138</b>	<b>\$474,864</b>	<b>\$12,726</b>

<sup>1</sup> Includes FTEs whose payroll obligations are supported by the NIH Common Fund.

**NATIONAL INSTITUTES OF HEALTH  
National Library of Medicine**

**Salaries and Expenses**  
(Dollars in Thousands)

OBJECT CLASSES	FY 2021 Enacted	FY 2022 President's Budget	FY 2022 +/- FY 2021
<b>Personnel Compensation</b>			
Full-Time Permanent (11.1)	\$42,248	\$43,209	\$961
Other Than Full-Time Permanent (11.3)	41,808	42,759	951
Other Personnel Compensation (11.5)	1,458	1,491	33
Military Personnel (11.7)	120	123	3
Special Personnel Services Payments (11.8)	1,888	1,931	43
<b>Subtotal Personnel Compensation (11.9)</b>	<b>\$87,522</b>	<b>\$89,514</b>	<b>\$1,992</b>
Civilian Personnel Benefits (12.1)	\$30,052	\$31,647	\$1,596
Military Personnel Benefits (12.2)	89	92	2
Benefits to Former Personnel (13.0)	0	0	0
<b>Subtotal Pay Costs</b>	<b>\$117,663</b>	<b>\$121,253</b>	<b>\$3,590</b>
Travel & Transportation of Persons (21.0)	\$420	\$428	\$8
Transportation of Things (22.0)	308	313	6
Rental Payments to Others (23.2)	140	143	3
Communications, Utilities & Misc. Charges (23.3)	477	486	9
Printing & Reproduction (24.0)	81	82	1
<b>Other Contractual Services:</b>			
Consultant Services (25.1)	87,601	89,351	1,750
Other Services (25.2)	70,790	73,563	2,773
Purchases from government accounts (25.3)	60,799	63,408	2,609
Operation & Maintenance of Facilities (25.4)	213	213	0
Operation & Maintenance of Equipment (25.7)	14,689	14,954	264
Subsistence & Support of Persons (25.8)	5	5	0
<b>Subtotal Other Contractual Services</b>	<b>\$234,098</b>	<b>\$241,494</b>	<b>\$7,397</b>
Supplies & Materials (26.0)	\$1,189	\$1,210	\$21
<b>Subtotal Non-Pay Costs</b>	<b>\$236,713</b>	<b>\$244,157</b>	<b>\$7,444</b>
<b>Total Administrative Costs</b>	<b>\$354,376</b>	<b>\$365,409</b>	<b>\$11,034</b>

**NATIONAL INSTITUTES OF HEALTH  
National Library of Medicine**

**Detail of Full-Time Equivalent Employment (FTE)**

OFFICE/DIVISION	FY 2020 Final			FY 2021 Enacted			FY 2022 President's Budget		
	Civilian	Military	Total	Civilian	Military	Total	Civilian	Military	Total
Division of Extramural Programs									
Direct:	18	-	18	21	-	21	21	-	21
Reimbursable:	-	-	-	-	-	-	-	-	-
Total:	18	-	18	21	-	21	21	-	21
Division of Library Operations									
Direct:	253	-	253	284	-	284	284	-	284
Division of Library Operations									
Reimbursable:	-	-	-	-	-	-	-	-	-
Division of Library Operations									
Total:	253	-	253	284	-	284	284	-	284
Lister Hill National Center for Biomedical Communications									
Direct:	34	-	34	38	-	38	38	-	38
Reimbursable:	-	-	-	-	-	-	-	-	-
Total:	34	-	34	38	-	38	38	-	38
National Center for Biotechnology Information									
Direct:	268	1	269	316	1	317	316	1	317
Reimbursable:	1	-	1	1	-	1	1	-	1
Total:	269	1	270	317	1	318	317	1	318
Office of the Director/Administration									
Direct:	59	-	59	67	-	67	67	-	67
Reimbursable:	13	-	13	13	-	13	13	-	13
Total:	72	-	72	80	-	80	80	-	80
<b>Total</b>	<b>646</b>	<b>1</b>	<b>647</b>	<b>740</b>	<b>1</b>	<b>741</b>	<b>740</b>	<b>1</b>	<b>741</b>
Includes FTEs whose payroll obligations are supported by the NIH Common Fund.									
FTEs supported by funds from Cooperative Research and Development Agreements.	0	0	0	0	0	0	0	0	0
<b>FISCAL YEAR</b>	<b>Average GS Grade</b>								
2018	11.9								
2019	11.8								
2020	11.8								
2021	11.8								
2022	11.8								

**NATIONAL INSTITUTES OF HEALTH  
National Library of Medicine**

**Detail of Positions<sup>1</sup>**

GRADE	FY 2020 Final	FY 2021 Enacted	FY 2022 President's Budget
Total, ES Positions	4	4	4
Total, ES Salary	789,200	797,092	815,186
General Schedule			
GM/GS-15	17	21	21
GM/GS-14	51	54	54
GM/GS-13	119	135	135
GS-12	109	113	113
GS-11	24	27	27
GS-10	0	0	0
GS-9	30	33	33
GS-8	30	30	30
GS-7	4	4	4
GS-6	1	1	1
GS-5	1	1	1
GS-4	3	3	3
GS-3	0	0	0
GS-2	4	4	4
GS-1	1	1	1
Subtotal	394	427	427
Commissioned Corps (42 U.S.C. 207)			
Assistant Surgeon General	0	0	0
Director Grade	0	0	0
Senior Grade	1	1	1
Full Grade	0	0	0
Senior Assistant Grade	0	0	0
Assistant Grade	0	0	0
Subtotal	1	1	1
Ungraded	272	309	309
Total permanent positions	392	488	488
Total positions, end of year	671	741	741
Total full-time equivalent (FTE) employment, end of year	647	741	741
Average ES salary	197,300	199,273	203,796
Average GM/GS grade	11.8	11.8	11.8
Average GM/GS salary	110,371	111,475	114,005

<sup>1</sup> Includes FTEs whose payroll obligations are supported by the NIH Common Fund.