NATIONAL INSTITUTES OF HEALTH
NATIONAL LIBRARY OF MEDICINE
PROGRAMS AND SERVICES FY 2005

TRANSPORTING THE WOUNDED.

U.S. DEPARTMENT OF HEALTH & HUMAN SERVICES

Further information about the programs described in this administrative report is available from the:

Cover: "Transporting the Wounded." A steel engraving (16 x 21 cm) in the NLM collection. By C.E. Wagstaff and J. Andrews after a drawing by Samuel Eastman ca. 185? Two Indians are transporting a sick or wounded comrade on a stretcher. NLM has an active program of support and outreach for American Indians, Alaska Natives, and Native Hawaiians (see pages 3–4).

NATIONAL INSTITUTES OF HEALTH

National Library of Medicine

Programs and Services

Fiscal Year 2005

U.S. Department of Health and Human Services
Public Health Service
Bethesda, Maryland

CONTENTS

Appendixes

Tables

PREFACE

The National Library of Medicine made notable progress on a number of fronts in Fiscal Year 2005. These accomplishments provide proof that the Library works at the crossroads of basic science, health care, public health, and informatics. NLM's programs frequently involve other components of the National Institutes of Health, other Federal agencies, and state and local governmental entities. Biomedical science, and the communication of advances resulting from it, is increasingly an interdisciplinary and interagency affair. Among the highlights:

- The Library embarked on an ambitious long-range planning initiative to ensure that it is optimally positioned in the coming years to serve the scientific community and the public.
- The "Go Local" feature of MedlinePlus, which connects users to local health organizations and services, continues to expand—four states were added this year.
- Hurricane Katrina prompted quick response from the NLM and from the National Network of Libraries of Medicine to create hurricane-related information sources and aid medical libraries in the devastated areas. NLM's Specialized Information Services provided handheld Wireless Information System for Emergency Responders devices (WISER) and the National Center for Biotechnology Information assisted in identifying victims using the latest DNA techniques.
- The National Center for Biotechnology Information's databases continued to expand: GenBank reached the milestone of 100 gigabases thanks to cooperation among international sequence repositories; a new Influenza Virus Resource was created in partnership with the National Institute of Allergy and Infectious Diseases to create the Influenza Virus Resource; and the PubChem database, an important component of the NIH Roadmap initiative, integrated compound data from more than 30 external sources, as well as the bioassay data that is beginning to be generated by high throughput screening centers.
- ClinicalTrials.gov, maintained by the Lister Hill Center, expanded its inclusion of international trials and the range of information it contains; 9,000 new trials were added in the last five months of the Fiscal Year.
- NLM is playing a key role in implementing the NIH Public Access Policy to ensure that taxpayer-funded biomedical research is widely and freely available.
- The Robert Wood Johnson Foundation joined with NLM's Extramural Programs in a new $3.6 million grant program to establish public health informatics tracks at four of the existing NLM-funded informatics training centers.

These are but a few of the many accomplishments of the Library in the past year. They are made possible through the efforts of a superb staff guided by more than eighty members of six formally constituted advisory committees. To all of them I offer my heartfelt thanks for a job well done.

_____

Donald A.B. Lindberg, M.D.
Director

OFFICE OF HEALTH INFORMATION PROGRAMS DEVELOPMENT

Elliot R. Siegel, Ph.D.
Associate Director

The Office of Health Information Programs Development (OHIPD) is responsible for three major functions:
- planning, developing, and evaluating a nationwide NLM outreach and consumer health program to improve access to NLM information services by all, including minority, rural, and other underserved populations;
- conducting NLM's international programs; and
- establishing, planning, and implementing the NLM Long Range Plan and related planning and analysis activities.

Planning and Analysis

The NLM Long Range Plan remains at the heart of NLM's planning and budget activities. Its goals form the basis for NLM operating budgets each year. All of the NLM Long Range Plan documents are available on the NLM Web site.

The Library is engaged in the process of reformulating its Long Range Plan, in this case for the period 2006–2016. The Long Range Plan for 1985–2005 foresaw and guided major developments at NLM, including articulating the need for the National Center for Biotechnology Information, the Visible Humans, the Unified Medical Language System, and programs of health information to the consumer.

An invitational Strategic Visions Working Group was convened by the NLM Board of Regents Subcommittee on Planning on April 11–12, 2005 in Washington, DC, to provide the broadest view of NLM's mission, current situation, and its potential future contributions to the health and wellbeing of America. Eighteen outstanding leaders from the health, business, and the library and information technology communities gave generously of their time and unique perspectives. The Working Group's vision incorporates a broad view of NLM's mission, current situation, and its potential future contributions to the health and well being of America to frame our considerations.

Based on the strategic vision articulated by the Working Group, four panels are developing specific recommendations for NLM:

Panel 1. NLM Resources and Infrastructure for the 21st Century
Panel 2. NLM Health Information for Underserved and Diverse Populations in the 21st Century
Panel 3. NLM Support for Clinical and Public Health Systems of the 21st Century
Panel 4. NLM Support for Genomic Science in the 21st Century

These panels are meeting in early FY 2006, with a final Long Range Plan to be prepared later in the year.

Based on the Long Range Plan, OHIPD documents NLM's progress in achieving its goals for a variety of purposes, including the Government Performance and Results Act and appropriations hearings, as well as NLM's involvement in a variety of disease and policy-related areas.

In addition to specific outreach and consumer health projects outlined below, OHIPD has overall responsibility for developing and coordinating the NLM Health Disparities Plan. This plan outlines NLM strategies and activities undertaken in support of NIH efforts to understand and eliminate health disparities between minority and majority populations. NLM's Health Disparities Plan is available on the NLM Web site.

Outreach and Consumer Health

NLM carries out a diverse set of activities directed at building awareness and use of its products and services by health professionals in general and by particular communities of interest. Considerable emphasis has been placed on reducing health disparities by targeting health professionals who serve rural and inner-city areas. Additionally, starting in 1998, NLM has undertaken new initiatives specifically devoted to addressing the health information needs of the public. These projects build on long experience with addressing the needs of health professionals and on targeted efforts aimed at making consumers aware of medical resources, particularly in the HIV/AIDS area.

NLM Coordinating Committee on Outreach, Consumer Health and Health Disparities

This office has convened and is chairing the NLM Coordinating Committee on Outreach, Consumer Health and Health Disparities (OCHD). This Committee plans, develops, and coordinates NLM outreach and consumer health activities. A number of the activities described below are conducted under the auspices of the OCHD.

American College of Physicians Physician Information Prescription Project

Doctors often prescribe medication after seeing a patient. But what if that doctor also wants to direct the patient to up-to-date, reliable, consumer-friendly information about a health concern? The American College of Physicians Foundation teamed with the NLM in FY 2003 to create the "Health Information Prescription" program. Doctors in several pilot states were given customized prescription pads that they can use to point patients to first-rate online health information in NLM's MedlinePlus database.

The Information Rx project was launched nationally on April 22, 2004, the opening day of the American College of Physicians (ACP) Annual Session in New Orleans. The joint project has been tested in Georgia and Iowa by more than 500 ACP internists and their patients. Among a variety of feedback tools yielding important findings, pre- and post-tests found that 97 percent of the participating internists made information referrals, with 59 percent using the prescription pads for information provided by ACP Foundation and NLM. Twenty percent of participating physicians also reported an increase in patients bringing Internet information to the office visit.

Internists who participated in the pilot programs said that MedlinePlus empowers patients (54 percent), explains difficult concepts and procedures (43 percent), and improves patient-physician communication (42 percent). The project was modified for the third stage of the pilot program in Virginia in March 2004 to partner with Virginia librarians as an additional resource to help patients use MedlinePlus. In 2005, the program was expanded to include a new collaboration with the Fisher Center for Alzheimer's Research to provide a disease-specific focus to attract new physician and patient participation. Pilots were launched in several Florida counties, and with the assistance of the American Medical Association Foundation, will be expanded statewide.

Consumer Health Diabetes Projects

NLM is actively exploring the use of new information technologies to enable diabetes patients to manage their disease and avoid or delay the onset of costly and debilitating complications, especially patients from minority and medically underserved populations.
In particular, we seek to learn how the use of NLM's MedlinePlus Web site, and other computer-based health information resources, can be helpful to patients, their families, and members of the public to learn about and understand the latest research news on diabetes, nutritional requirements, tests, devices, secondary prevention techniques, and for obtaining answers to patient-specific questions. In the clinical setting, the principal hypothesis is that MedlinePlus can reinforce and supplement the information provided by physicians, nurses and health educators. A related hypothesis is that a combination of individualized training and access to publicly available computer resources at hospital libraries and elsewhere in the community can help reduce the health disparities experienced by minority populations that have less ready access to computer-based health information in the home, school and workplace than the majority population.

The goal is to develop, design, implement and evaluate a comprehensive program of diabetes-focused outreach initiatives in collaboration with academic health science centers and libraries, clinical centers, community-based organizations and voluntary health organizations.

The latest initiative in collaboration with the Upper Cardozo Health Center in Washington, D.C. and George Washington University, undertakes a controlled field experiment with patients enrolled in the Diabetes Health Disparities Collaborative wherein diabetes patients will receive an individualized information technology-based intervention to complement their regular patient education program. Education and training will be given to (1) physicians on how to incorporate the Information Prescription Project into their daily office practice routine; (2) medical assistants to show patients how to use MedlinePlus in English and Spanish; and (3) clerks on how to instruct patients to use MedlinePlus in waiting rooms during clinic hours. A wealth of routinely collected clinical patient data (e.g., cholesterol) will give NLM a unique opportunity to provide objective evidence of the impact of MedlinePlus use on patient health outcomes.

Web Evaluation

The Internet and World Wide Web now play a dominant role in dissemination of NLM information services. And the Web environment in which NLM operates is rapidly changing and intensely competitive. These two factors combined suggested the need for a more comprehensive and dynamic NLM Web planning and evaluation process. The continuing Web evaluation priorities of the OCHD include: (1) quantitative and qualitative metrics of Web usage; and (2) measures of customer perception and use of NLM Web sites. During FY 2005, the OCHD continued to pursue an integrated approach intended to encourage exchange of information and learning within NLM, and help better inform NLM management decision-making on Web site research, development, and implementation. The year's evaluation activities included: access to a syndicated telephone survey of the U.S. public's online and offline health information seeking behavior; analysis of NLM Web site log data;

and access to Internet audience measurement estimates based on Web usage by user panels organized by private sector companies. The OCHD Committee and OHIPD continue to explore and test a range of internal and external Web evaluation methods and applications.

Also during FY 2005, OHIPD continued to collaborate with other units of NIH to implement a trans-NIH online user survey project based on the American Customer Satisfaction Index (ACSI), with significant initial and supplemental funding support from the NIH Office of Evaluation. The project has extended the ACSI online user survey methodology to about 60 NIH Web sites at 28 NIH components, with expected benefits to both specific Institute Web site development and to trans-NIH collaboration on better understanding and improving the overall NIH Web presence. The project includes multi-level evaluation objectives: (1) strengthening of each participating Institute's Web evaluation capability and activity; (2) sharing of Web evaluation learning and experience on a trans-NIH basis; (3) aggregation of ACSI results and learning on a trans-NIH basis; and (4) sponsorship of an NIH-wide staff workshop that will highlight the contributions and challenges of the ACSI from the NIH perspective, consolidate lessons learned, and identify future directions and opportunities. The project is being managed by a trans-NIH Survey Leadership Team, of which NLM is a member. The primary ACSI contractor is ForeseeResults Inc. via an Interagency Agreement between NLM and the Federal Consulting Group/U.S. Department of the Treasury. The primary evaluation contractor is Westat Inc., which is conducting evaluation activities under the guidance of and in coordination with the ACSI Leadership Team and participating NIH units.

Tribal Connections

NLM has continued to focus on improving Internet connectivity and access to health information services in American Indian and Alaskan Native communities. Phase I (Pacific Northwest) and Phase 2 (Pacific Southwest) of tribal connections are complete, with a final project evaluation published in Wood et al., Tribal Connections Health Information Outreach: Results, Evaluation, and Challenges, Journal of the Medical Library Association, Vol. 91, January 2003, pp. 57-66. Also, Phase 3 is complete. In Phase 3, more intensive community-based outreach and training were implemented at select Phase 1 and 2 sites to assess if these community-based approaches significantly enhance the project impacts on health information, behavior, and outcomes. A Phase 3 evaluation report is available from the Pacific Northwest Regional Medical Library, University of Washington, Seattle.

NLM has funded a Phase 4, in collaboration with the University of Utah (Midcontinental Regional Medical Library), emphasizing the development of Web-based tribal health information resources in the Four Corners Region (AZ, CO, NM, UT). The Phase 4 was completed in FY 2005, and an evaluation report is available from the Midcontinental RML. NLM has also funded a Phase 5 that will include outreach to Four Corners public libraries serving American Indians in that region, and the convening of an NN/LM meeting to consolidate lessons learned from Native American outreach to date.

Other Native American Outreach

In 2005 NLM/OHIPD again partnered with NIH and NLM Equal Employment Opportunity offices to participate in the NIH American Indian Pow-Wow Initiative. This included exhibiting at nine pow-wows mostly in the Mid-Atlantic area. Also NLM/OHIPD participated with NIH in the National Museum of the American Indian Powwow in Washington, D.C. An estimated 15,000 persons visited the NLM booth over the course of these pow-wows. These activities proved to be another viable way to bring NLM's health information to the attention to segments of the Native American community and the general public.

In other parts of the country, in FY 2005 OHIPD supported several projects in the Dakotas, Hawaii, and Alaska. These projects resulted largely from the Native American Listening Circles conducted in FY 2003–2004:

- North Dakota—University of North Dakota Health Sciences Library (via the Greater Midwest RML), phase II of health information outreach to North Dakota tribal colleges and tribal health;
- South Dakota—University of South Dakota Health Sciences Library (via the Greater Midwest RML), project to extend health information outreach to South Dakota tribal colleges and health;
- North Dakota—Cankdeska Cikana Community College (via the Greater Midwest RML), Spirit Lake Nation, Ft. Totten, ND, project to develop a health-related educational program at the Community College, and associated improvements in the technical and information resource infrastructure at the tribal library;
- North Dakota—MHA Systems Inc., a tribal enterprise of the MHA Nation, economic development outreach project to provide outreach assistance to a tribal information technology company that would ultimately result in jobs creation on the reservation (in this case, the Ft. Berthold Indian Reservation, home to the MHA Nation); the project is intended to improve the competitive capabilities and posture of MHA

Systems Inc., and also to refine, test, and strengthen the company's core scanning services (through a pilot test of scanning materials provided by NLM);

- Hawaii—Papa Ola Lokahi (via the Pacific Southwest RML), two Native Hawaiian Community Health Education Projects—Community of Miloli'i, Hawaii (The Big Island)—intended to increase the knowledge of community members about available health information and health resources by providing computer hardware and software to the community's library and training for the librarian and other community members, by increasing multi-media resources at the Miloli'i Community Library, and by supporting community-based initiatives which have their foundation in the Hawaiian concepts of health (which involves a balance between body, mind and spirit);

- Hawaii—Waimanalo Health Center, Oahu (Windward side)—intended to increase the knowledge of community members about health information resources in order to better understand their own health conditions or health conditions of family members and to enable more effective self-management and more informed communication with health service providers, to be achieved by providing training and access to Web-based sources of health and medical information; and

- Alaska—Buckland Alaska Native Village (via Pacific Northwest RML), a project to improve the health information infrastructure at the village school and/or health clinic, including needs assessment, technical upgrades, and associated training and outreach.

Outreach to Hispanics

The Lower Rio Grande Valley Hispanic Outreach Project was a collaboration with the University of Texas at San Antonio Health Sciences Center to conduct a needs assessment and various health information outreach projects with Hispanic-serving community, health, and educational institutions. The initial Lower Rio Grande Project is complete and, based on the project results, NLM has funded a series of follow-on projects focusing on outreach to Hispanic populations in South Texas. One completed follow-on project involves Hispanic residents of the Lower Rio Grande Valley who live in colonias. This project involves collaboration with Texas A&M University as well as the University of Texas at San Antonio, and its Regional Academic Health Center in Harlingen, TX. Also, OHIPD funded two projects building on the very successful pilot project at MedHigh in the Lower Rio Grande Valley,

where high school peer tutors were trained and then in turn taught their peers about health information resources available from NLM. The pilot project received several awards for outstanding performance and service to the students and surrounding communities. The follow-on projects extend the MedHigh concept and collaborations to other magnet high schools in the Lower Rio Grande Valley and prospectively in other geographic areas.

Community-Based Outreach Symposium

On December 2–3, 2004, NLM sponsored a Symposium on Community-Based Health Information Outreach in collaboration with the National Networks of Libraries of Medicine. The symposium was co-chaired by Wayne Peay, Director, Eccles Health Sciences Library, University of Utah, and Maxine Rockoff, Director, Division of Information Management, New York Academy of Medicine. The principal goal of the Symposium was to explore new models of health information outreach that are emerging as technology dramatically changes the abilities of medical and health sciences libraries to provide resources and services beyond their traditional institutional boundaries. The Symposium emphasized consumer health information outreach working through community-based organizations. NLM's Strategic Plan to Reduce Health Disparities provided an introductory framework for the Symposium, and the Plan was reviewed in a post-Symposium work session with special emphasis on Native American Outreach.

Symposium speakers identified lessons learned from previous NLM and NN/LM outreach projects; discussed the differing perspectives of those engaged in outreach at the grass-roots level; highlighted alternative evaluation models and criteria for judging "success"; and contrasted the theoretical underpinnings of NLM outreach that have a basis in both communication research and library/information sciences. The formal program ended with a futurist's perspective on societal and medical influences. However, the true conclusion picked up where the program had begun, with an affirmation of former Board of Regents chair, Eugenie Prime's keynote challenge that NLM pursue outreach with passion, remain unfearful of "promiscuous partnerships," and embrace a commitment to execution—the discipline of getting things done.

Several white papers and 17 poster abstracts are available at the Symposium web site, http://medstat.med.utah.edu/symposium. The white papers formed the basis of a proceedings, which were published as a special supplement issue of the *Journal of the Medical Library Association,* Vol. 93, No. 4, October 2005.

International Programs

The OHIPD projects described in this section are funded through support leveraged within NLM as well as with outside agencies.

MIMCom: A Malaria Research Network for Africa

Since 1997, NIH has led an international effort to provide malaria researchers in Africa with full access to the Internet and the resources of the World Wide Web. This project began with NIH's leadership in the Multilateral Initiative on Malaria (MIM) in which African scientists identified electronic communication and access to scientific information as critical in the fight against the devastating and economically debilitating effects of malaria in developing countries.

As a part of MIM, NLM, working in partnership with organizations in Africa, the United States, the United Kingdom and Europe, has created MIMCom.Net, the first electronic malaria research network in the world. Using satellite technology, the network provides African scientists with full access to the Internet and the resources of the World Wide Web, as well as access to current medical literature. The African research sites are of recognized high quality, require improved communications to accomplish ongoing research, and have the necessary resources to purchase equipment and sustain the system.

The Web site, http://www.nlm.nih.gov/mimcom, has links to MEDLINE, a variety of free online journals, databases, malaria-related sites, and general information. An NLM reference librarian serves as the Webmaster and is expanding the site to include special news releases and articles of interest to researchers.

MIMCom has evolved in order to support the shifting needs of the malaria research sites in Africa. In 1998, the network started with a microwave link to the Internet in Bamako, Mali, and has since assisted 19 other sites in 12 countries. In the intervening years, the telecommunications revolution has moved forward—technology has changed, along with the Internet itself, the latter now bringing us spam and viruses undreamt of in the not too distant past. Conversely, where there was once little or nothing in terms of telecommunications options, there are now, in some but not all instances, a number of players providing useful services, resulting in competitive pricing. Additionally, some sites have experienced dramatic growth and are no longer properly served by the system as it currently exists.

The next very practical phase for MIMCom is to focus on medical informatics. We have begun to do this with the Antimalarial Drug Resistance Network. This network, which has set up a secure server so that scientists can share raw data and post data summaries,

addresses the need to help scientists make innovative use of their new research tools to facilitate new ways of working together. The article "Tying up lions: multilateral initiative on malaria communications: the first chapter of a malaria research network in Africa" (http://www.ajtmh.org/cgi/content/full/71/2_suppl/259) provides more information about MIMCom.

MIMCom Evaluation

NLM has also studied the contribution of enhanced connectivity to the performance of malaria research staff working in remote sites in Africa. The main objectives were to examine the use of MIMCom supported information technology (IT) by scientists, students, and administrative personnel and to determine what contribution this intervention made to their professional performance. Four sources of data were used for this assessment: (1) A self-administered online questionnaire directed to biomedical scientists, students, and support staff via the MIMCom Website to MIMCom users from Cameroon, Gabon, Ghana, Kenya, Malawi, Tanzania, and Uganda; (2) MIMCom summary statistics on bandwidth use and the type of Websites visited by the study sites; (3) Full text document requests via the MIMCom website; and (4) Information from site directors regarding publication of journal articles.

MIMCom makes possible enhanced access to the Internet and to medical literature, thus facilitating communication, information retrieval, document delivery, collaboration on proposals, and preparation of scientific papers for publication. By providing full access to the Internet and the resources of the Web, MIMCom has been shown to be invaluable to malaria researchers and their institutes in Africa. This access has increased visibility of scientists in their respective institutions and provided opportunities for stronger engagement with the international scientific community. A more detailed description of this study is available in: Royall, J., Bennett, M., van Schayk, I., Kamau, N., Alilio, M. (2005), Crossing the digital divide: the contribution of information technology to the professional performance of malaria researchers in Africa. *African Health Sciences*, 5, 3, 246–254.

MIMCom: New Second Phase

NLM has provided the leadership, resources, and glue for this first phase of MIMCom and has achieved the goals of the initial mandate. With generous support from the Swedish International Development Agency (SIDA), more sites are now able to benefit from improved access to medical literature and to the Internet through MIMCom. The goal is to promote further capacity building, resulting in strong African research

and leadership characterized by networks and sharing of information.

As evident in this project, access to Internet and medical literature alone will not make the difference in stronger research and ultimately better health in Africa. Funding, people, partnerships, and sustainability are all critical pieces of the bigger picture. Some reside inside of NLM and others on the outside. Some current examples: the new MIMCom sites (six more thus far) are being funding by SIDA; the Webmaster and Website by NLM; server and Website for Antimalarial Drug Resistance Network by NLM with technical support by SIDA, the World Bank, and the World Health Organization, and the MIMCom weekly newsletter funded by NLM/OHIPD.

In partnership with NLM, the MIM Secretariat in Stockholm, the SIDA, and a team at the Royal Institute of Technology in Stockholm have engaged students from Europe and developing countries in working together to design and implement IT solutions in Africa. As a result, connectivity now enhances research activities at two new MIMCom sites, in Mozambique and Uganda, with initial site visits completed for three additional installations in Nigeria, Zambia, and the Congo.

The newsletter— MIMCom Malaria News Update—is a weekly publication sent to a listserv comprising researchers at 20 sites. This e-mail alert service provides regular malaria news updates to support malaria research in Africa and networking around the world and has over 1,000 subscribers.

African Medical Journal Editors Partnership Project

The African Medical Journal Editors Partnership Project is a collaborative project with NIH's Fogarty International Center. The objective is to create four partnerships between four African medical journals and journals from the U.S. and U.K. for the purpose of strengthening the African journals. NLM's specific objective is strengthening African journals so that they are able to get into MEDLINE, and, as a result, make African research available to the world.

Partnerships have been initiated between editors of the following journals: *Ghana Medical Journal* and *The Lancet*; *Malawi Medical Journal* and the *Journal of the American Medical Association*; *African Health Sciences* and *British Medical Journal*; and the *Mali Medical Journal* and the *American Journal of Public Health* and the *Environmental Health Perspectives/National Institute of Environment and Health Sciences journal*.

Partnerships can encompass sharpening of business, editorial, and technical aspects of: editorial skill development; training for authors, reviewers, and editorial board; sharing manuscripts; joint commissioning of articles; exchange of editorial content; staff exchange for skills and experience sharing; increased publication of local research; and a survey of a journal's target audience.

Training and On-the-Ground Development

NLM continues to be active in the training of medical librarians, including programs in which the trainers train others.

- NLM participates in the biannual Association for Health Information and Libraries in Africa conference by offering workshops in PubMed training and sponsoring the travel of African librarians associated with the MIMCom project.
- Librarians from Vietnam have come to NLM for training in indexing and MeSH, so they could begin to make their own collections available to physicians and health workers in that country. Additional visits by Vietnamese librarians are planned.
- NLM's Associate Program has an international fellow who returns home with expertise and resources to carry out projects locally. Fellows from Africa have included librarians from Kenya, Zambia, and Mozambique.
- The partnership for training and information access at Faculty of Medicine, Makerere University, Uganda, provides training in searching of NLM databases for medical students and faculty. The Albert Cook Library at the Medical School is also a testing site for PubMed on Tap for use on campus and in the field.
- The MedlinePlus for Tropical Regions is a new project that has already put together two prototype interactive MedlinePlus tutorials featuring tropical disease issues. They were created and field tested by African doctors and scientists at the Faculty of Medicine, Makerere University. One outcome of this program is that more African scientists are joining in the international community. Their eloquent voices and moving stories are on the Web at http://www.ajph.org/cgi/content/full/95/4/559.

Global Internet Connectivity

End-to-end performance of the Internet, on both national and global scales, continues to be important to NLM. This includes the further exploration of the methods and metrics needed to better understand the quality of Internet performance from the end user perspective. NLM, which is a leader in this field, has implemented Phases I and II of its own Internet connectivity performance monitoring network, starting with select U.S. sites (the eight Regional Medical Libraries and several Resource Libraries) but envisioned eventually to extend to other U.S. sites and some international sites. A "net-centric" planning and evaluation workshop is scheduled for FY 2006. The network monitoring is moving from experimental to operational status, as the performance data are important to NLM's own network management.

International MEDLARS Centers

Bilateral agreements between the Library and 18 public institutions in foreign countries allow them to serve as International MEDLARS Centers. As such, they assist health professionals in accessing MEDLINE and other NLM databases, offer search training, provide document delivery, and perform other functions as biomedical information resource centers. These Centers are in Canada, China, Egypt, France, Germany, Hong Kong, India, Israel, Italy, Japan, Korea, Mexico, Norway, Russia, South Africa, Sweden, United Kingdom, and the Pan American Health Organization (BIREME) in Brazil.

International Visitors

In FY2005, the Office of Communications and Public Liaison and the History of Medicine Division arranged for 358 tours—106 regular daily (1:30 p.m.) tours and 252 specially arranged tours. There were 7408 visitors in all. They came from the following 48 countries:

> Argentina, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Bolivia, Botswana, Brazil, Canada, Chile, China, Colombia, Czech Republic, El Salvador, England, France, Georgia, Ghana, Greece, Haiti, Hungary, India, Israel, Japan, Kazakhstan, Kenya, Korea, Mali, Mexico, Moldova, Panama, Paraguay, Peru, Russia, South Africa, Spain, Sri Lanka, St. Lucia, Switzerland, Taiwan, Tanzania, Thailand, Trinidad, Uganda, Ukraine, United States, Zambia.

LIBRARY OPERATIONS

Becky J. Lyon
Deputy Associate Director

The Library Operations (LO) Division is responsible for the basic services that ensure access to the published record of biomedical science and the health professions. LO acquires, organizes, and preserves NLM's comprehensive collection of biomedical literature; creates and disseminates controlled vocabularies and a library classification scheme; produces authoritative indexing and cataloging records; builds and distributes bibliographic, directory, and full-text databases; provides back-up document delivery, reference service and research assistance for the nation; helps varied user groups to make effective use of NLM products and services, and coordinates the National Network of Libraries of Medicine to improve access to health information services across the United States. These services provide an essential foundation for NLM's outreach programs to health professionals and the general public. They also support the Library's focused programs in AIDS, health services research, molecular biology, and toxicology and environmental health.

In addition to its basic services, LO develops and mounts major historical exhibitions; carries out an active research program in the history of medicine; works with other NLM program areas to enhance NLM products and services; conducts research related to current operations and services; directs and sponsors training programs for health sciences librarians; and contributes to the development of national health data standards policy and to the production and dissemination of clinical vocabulary standards.

LO employs a multidisciplinary staff of librarians, technical information specialists, subject experts, health professionals, historians, museum professionals, and technical and administrative support personnel and relies on the services of a wide range of contractors. LO is organized into four major Divisions: Bibliographic Services (BSD), Public Services (PSD), Technical Services (TSD), and History of Medicine (HMD); three units: the Medical Subject Headings (MeSH) Section, the National Network of Libraries of Medicine Office (NN/LM), and the National Center on Health Services Research and Health Care Technology (NICHSR); and a small administrative office. LO staff members participate actively in efforts to improve the quality of work life at NLM, including the Diversity Council and the NLM Intranet.

Most LO activities are critically dependent on automated systems developed and maintained by NLM's Office of Computer and Communications Systems (OCCS), National Center for Biotechnology Information (NCBI), or Lister Hill National Center for Biomedical Communications (LHC). LO staff work closely with these program areas on the design, development, and testing of new system features.

Program Planning and Management

Priorities for LO programs are based upon the goals and objectives identified in the NLM Long Range Plan 2000–2005 and, where appropriate, in the NLM Strategic Plan to Reduce Racial and Ethnic Disparities. As described in the previous chapter, "Office of Health Information Programs Development," an NLM Long Range Plan for 2006–2016 is currently being developed under the auspices of the Library's Board of Regents.

The current NLM Long Range Plan includes a strong emphasis on the influence of the rapidly evolving electronic publishing field and the increasing role of the Internet in improving access to health information. In FY 2005 a major effort in this area involved working with the NCBI and the Office of the Director to implement the NIH Public Access Policy.

LO staff also began to discuss the substantial challenges that electronic publishing poses to NLM's ability meet its statutory responsibility to collect, preserve, and provide access to the biomedical literature. In FY 2005, LO continued to review and revise policies, procedures, and services to reflect shifting workloads; to use electronic information to enhance basic operations and services; and to work with other NLM program areas to ensure permanent access to electronic information. LO also focused considerable attention on working with other NLM program areas to meet the Library's expanded responsibility for distribution of standard clinical vocabularies within the UMLS Metathesaurus. Significant progress has been made in transitioning responsibility for the Metathesaurus production system from LHC to LO and OCCS, with work to be completed by the end of calendar 2006.

Although many of its efforts are directed toward creating and promoting use of electronic information resources and to supporting NLM's high priority outreach initiatives, LO also devotes substantial resources and attention to the care and handling of NLM's extensive collections of physical library materials and to the space and environment in which staff and patrons work and collections are stored. In FY 2005, video cameras were installed in the Main and History of Medicine Reading Rooms to provide improved security for the NLM collections. LO also continued to make improvements to conditions in the existing NLM library building and to develop strategies for handling the projected growth of the collections until a new facility becomes available. These strategies are discussed elsewhere in this report.

The Request for Proposals (RFP) for new 5-year contracts for eight Regional Medical Libraries in National Network of Libraries of Medicine was issued on April 15, 2005. The statement of work for 2006–2011 continues to emphasize outreach to the public, network members, and health professionals with a focus on minority and underserved populations. The contracts also seek to build and improve collaborations with community-based organizations as an effective means of reaching these populations. In addition to the basic contract, the solicitation invited proposals for the National Training Center and Clearinghouse, the Outreach Evaluation Center which will include outreach mapping, and a Web Services and Technology Operations Center. Proposals were received in July and initial technical reviews were conducted in September and October by three review teams, each consisting of an academic health sciences librarian, a hospital librarian, a health professional, and an NLM staff member. Questions identified during these reviews have been sent to each offeror. The process of selecting the successful bidders and awarding the new NN/LM contracts will be completed in the spring of 2006, after site visits in some regions.

In FY 2005, LO's Administrative Office continued to assist managers, supervisors and staff with the transition to a range of new administrative systems and the challenges of changes in the functioning of human resources within NIH. LO continues to encourage its staff to take advantage of flexiplace work arrangements as appropriate. More than 84 employees work at home at least one day per week.

Collection Development and Management

NLM's comprehensive collection of biomedical literature is the foundation for many of the Library's services. LO ensures that this collection meets the needs of current and future users by updating NLM's literature selection policy; acquiring and processing relevant literature in all languages and formats; organizing and maintaining the collection to facilitate current use; and preserving it for subsequent generations. At the end of FY 2005, the NLM collection contained more than two and a half million volumes and more than six million other physical items, including manuscripts, microforms, pictures, audiovisuals, and electronic media.

Selection

In FY 2005, selectors worked on a number of projects to enhance the NLM collections. Published works in the burgeoning field of nanobiotechnology and related fields were reviewed for addition to the collection. Works were identified and selected from Latin American countries with little representation in the collection. Orders and claims for Eastern European vendors were reviewed, potential vendors for Greek publications were identified and contacted, and relations were reestablished with the Library's vendor in Israel. NLM's Chinese monographic vendor went out of business and possible replacement vendors were identified. Selectors took several steps to identify published works to add to the collection in the area of Native Hawaiian health and healing. The Library is also reviewing and acquiring doctoral theses that contain original documentation of Native Hawaiian health and healing practices, policies, history, public health initiatives, and research.

Acquisitions

TSD received and processed 157,006 contemporary physical items (books, serial issues, audiovisuals, electronic media), which is slightly above last year's total. The increase in electronic publishing has not yet had a significant effect on the number of physical items that NLM acquires. A net total of 29,529 volumes and 587,474 other items (including nonprint media, manuscripts and pictures acquired by HMD) were added to the NLM collection. In FY 2005, 2,694 licensed and 251 free electronic journals were available to NLM users, many through the NIH Library licenses to Elsevier's Science Direct and Wiley Interscience. NLM purchased the Meridian electronic resource management system, which will be used to track the acquisition and licensing processes for electronic journals and databases.

HMD acquired many splendid early printed books, manuscripts, images, and historical films for the NLM collection in FY 2005. Important books acquired included: Conrad Schellig's *In Pustulas Malas Morbum* (Heidelberg, ca. 1496), which is perhaps the first German-language publication on syphilis; Rhazes (865–circa 925) *Tibb Almansorem*, an early seventeenth-century Arabic manuscript of his systematic treatise of medicine; a printed Tibetan work, *Rgyud-bhzi*, comprised of four treatises forming the basis of Tibetan medicine; and Thomas Willis' *Affectonum Quae Dicuntur Hystericae & Hypochondriacae Pathologia Spasmodica Vindicata...*(London, 1670), which argued that hysteria was a disease of the nervous system, not the uterus.

Additions to the archival and manuscript collections included the American Medical Association Card File of Deceased Physicians donated by The National Genealogical Society of Arlington, which contains biographical information for about 350,000 nineteenth- and twentieth-century physicians in the U.S. and Canada; the records of Sheppard-Pratt Hospital, a psychiatric hospital in Baltimore that stressed vocational

regimens for patients; additions to the Louis Sokoloff, Julius Axelrod, C. Everett Koop, Joshua Lederberg, and Adrian Kantrowitz collections; the papers of Lasker Award winners Edward Fries and Sol Spiegelman; and the papers of Karen Davis, President of the Commonwealth Foundation, who is noted for her work in health services research.

Important additions to the prints and photographs collection included another generous donation from William Helfand of 10,000 postcards depicting public health issues, and a collection of 2,000 nineteenth- and twentieth-century post cards depicting nursing practice in various countries. HMD negotiated a long-term plan to purchase a collection of nearly 8,000 items on Chinese public health assembled by Don Cohn. The materials from the People's Republic of China from the 1950s, 1960s, and 1970s are unique in western repositories.

NICHSR, HMD and LHC collaborated to expand the collection of interviews with eminent researchers as part of the effort to document the history of health services research. In addition, NICHSR arranged to receive a donation of microfiche from the Kaiser Permanente Center for Health Research. This collection covers 1964 to about 1985 and represents a very large health services research reference collection, including fugitive documents, article reprints, and even some early books.

LO chaired an NLM-wide working group to make recommendations on the acquisition of NIH videocasts. The group was expanded to include a representative from NIH's Center for Information Technology (CIT) and a joint NLM/CIT agreement resulted which calls for selection of materials from the NIH Videocasting site; improved titling, cataloging, and searching capabilities; support for better image quality and preservation; and voice-recognition keyword searching.

Preservation and Collection Management

LO undertakes a broad range of activities to preserve NLM's archival collection and keep it readily accessible for use. These activities include: binding, microfilming, conservation of rare and unique materials, book repair, maintenance of appropriate storage and environmental conditions, and disaster prevention and response. LO distributes data about what NLM has preserved to avoid duplicate effort by other libraries. LO works with other NLM program areas to conduct experiments with new preservation techniques as warranted and to promote the use of more permanent media and archival-friendly formats in new biomedical publications.

New purchase orders for microfilming, microfilm preparation, and audiovisual duplication were awarded. In FY 2005, LO bound 18,417 volumes,

microfilmed 1,564 volumes, repaired 2,095 items in NLM's onsite repair and conservation laboratory, made 936 preservation copies of motion pictures and videos, and conserved 140 rare items. The 40 percent decrease in the number of volumes microfilmed and the increase in the number of audiovisuals copied reflect the revision of NLM's preservation priorities that occurred in FY 2004. Planning for a digital preservation pilot project to scan nineteenth-century pamphlets on cholera was well under way in FY 2005.

Dwindling space for growth of the NLM collections continues to require significant resources for shifting of collections and contingency planning in the event that NLM does not receive funding for a new facility. Over the course of 167 years, the NLM has repeatedly outgrown its physical facilities as its collection grows. This unparalleled resource of books, journals, and other materials contains much that does not exist elsewhere and thus, as hospital and research libraries face increasing budget and space constraints, the NLM collection will assume even greater importance to the nation. Unfortunately, shelf space will soon be filled, and some special collections, such as modern manuscripts, historical audiovisuals, microfilms, prints, and photographs are already out of space.

During FY 2005, LO began a major project to weed the Z classed monographs and serials in the general collection in order to free up space for shelving processed modern manuscripts. The Z collection contains many works that are out of scope for the NLM collection. A purchase order for an inventory of the serial collection was awarded on September 30. This required a year of intensive planning involving several LO staff and assistance from the Office of Computer and Communications Systems and included sampling the collection, developing a statement of work, having wireless installed in the stacks, and obtaining specially designed carts to accommodate thin client computers that will operate in compact shelving. A major shift of serials published from 1990–1994, from space occupied on the B1 level of building 38 to the B3 level also began in FY 2005. When the shift of these serials is completed in 2006, space for serials on B3 will be completely filled. Current estimates are that onsite space for all collections will be filled by 2010. Staff is investigating a number of options for additional collection storage space including strengthening of floors in the existing building to support installation of compact shelving and the feasibility of installing high density shelving.

Permanent Access to Electronic Information

The preservation of electronic information presents unique challenges that are not yet fully understood. NLM's general approach to addressing these challenges

is to use NLM's own electronic services and publications as test-beds and to work with other national libraries, the National Archives and Records Administration, and other interested organizations to develop, test, and implement strategies and standards for ensuring permanent access to electronic information. LO works closely with other NLM program areas on activities related to the preservation of digital materials. LO chaired an NLM working group in FY 2005 charged with recommending to the Board of Regents areas for expansion of the Library's digital preservation research program. The Board approved the expansion of NLM's internal efforts to encompass collaborative research into interactive digital publications as well as NLM support for external digital preservation research and development, complementary to the Library of Congress National Digital Information Infrastructure and Preservation Program. The external program will include developing, testing, and implementing tools, standards and strategies for permanent access to electronic information and will also result in the preservation of important digital biomedical information.

PubMed Central, a digital archive of medical and life sciences literature developed by the National Center for Biotechnology Information, is NLM's primary test-bed for the development of procedures and methods for ensuring permanent access to electronic journals. LO continues to assist NCBI in expanding current deposits to PubMed Central by soliciting the participation of additional journals, primarily in the fields of clinical medicine, health policy, health services research, and public health. In addition, the Public Services Division also works closely with NCBI to scan and add the backfiles of journals, depositing current issues in the digital archive. The back issue scanning project proceeded smoothly throughout the year and is vastly increasing the material in PubMed Central. At the end of FY 2005, more than 287,000 articles have been made available. Complete runs of 20 journals have been processed and more than 70 additional journals are in production. Each journal is made available in its entirety, including front and back covers, tables of contents, administrative material such as masthead and editorial boards, and advertisements. Views include a full-text summary, HTML view, separate views of TIFF images, and the full PDF. Titles scanned through an agreement with the Wellcome Trust and the Joint Information Systems Committee in the United Kingdom which are nearing completion include *Annals of Surgery, Journal of Physiology, Biochemical Journal, and Journal of Anatomy*.

NLM is using its own main Web site as a test-bed for procedures and mechanisms for ensuring permanent access to electronic information published by government agencies and private non-profit institutions.

The Archives section of the site continues to be the repository for Web pages which have been designated as Permanent with unchanging or stable content and which have become outdated or replaced by newer versions. The archive will be expanded to incorporate other NLM Web sites.

Vocabulary Development and Standards

LO produces and maintains the Medical Subject Headings (MeSH), a subject thesaurus used by NLM and many other institutions to describe the subject content of the biomedical literature and other types of biomedical information; develops, supports, or licenses the U.S. vocabularies designed for use in patient records and clinical decision support systems; and works with the Lister Hill Center and OCCS to produce the Unified Medical Language System (UMLS) Metathesaurus, a large vocabulary database that incorporates many vocabularies, including MeSH and other vocabularies produced or supported by NLM. A multi-purpose knowledge source used in operational systems and informatics research, the Metathesaurus also serves as a common distribution vehicle for classification, code sets, and vocabularies designated as standards for U.S. health data.

LO represents NLM in federal initiatives to select and promote use of standard clinical vocabularies in patient records and administrative transactions governed by the Health Insurance Portability and Accountability Act of 1966 (HIPAA). In this capacity, LO staff members serve on the Department of Health and Human Services Data Standards Committee, provide staff support to the National Committee on Vital and Health Statistics (NCVHS) Standards and Security Subcommittee, participate in the Public Health Data Standards Consortium, and contribute to the Federal Health Architecture Consolidated Health Informatics Work Group. In FY 2004, in recognition of the Library's standards activities and expertise in health information technology, the Secretary of Health and Human Services (HHS) acted upon an NCVHS recommendation and designated NLM as the coordinating center for standard clinical terminologies. Funds were again transferred to NLM from other HHS agencies in FY 2005 to assist with these responsibilities.

The Secretary also selected NLM as the operational home of the Commission on Systemic Interoperability (CSI), which was established by the Medicare Modernization Act of 2003 to develop a comprehensive strategy for the adoption and implementation of health care information technology standards. LO provided administrative support to CSI throughout FY 2005. CSI is expected to release its report in October 2005. NLM staff is working with the Department of Health and Human Services, Office of

the National Coordinator for Health Information Technology to ensure NLM vocabulary standards efforts are fully coordinated with the efforts being led by the Office. This coordination included participation by LO staff in the technical review for the Standards Harmonization Process RFP and the Compliance Certification Process RFP issued by the Office.

Medical Subject Headings (MeSH)

The 2006 edition of MeSH contains 23,880 main headings, 136 publication types, 83 subheadings or qualifiers, and 154,371 supplementary records for chemicals and other substances. For the 2006 edition, the MeSH Section added 933 new descriptors, replaced 138 descriptors with more up-to-date terminology, and deleted 50 descriptors.

The 2006 vocabulary reflects work to reorganize and update the vocabulary for vitamins, expand terminology for membrane transport modulators including channel blockers, revise terminology for endocrine system diseases, and add new vocabulary related to adipose tissue, arteriosclerosis, obesity, bariatric surgery, and metabolism. In addition to continuing to ensure that taxonomies of organisms remain current with recognized authorities, the MeSH section completed the revision of the proteins, and continued work on the pharmacologic actions of drugs. A new tree, publication characteristics, was created for the publication types and for the research support tags. Three very infrequently searched Publication Types (Review of Reported Cases; Review, Multicase; and Review, Tutorial) have been deleted for 2006 MeSH and added as entry terms (see references) to the more general Review [Publication Type].

MeSH is translated into many other languages by organizations around the world, including a number of NLM's international MEDLARS partners. The Web-based MeSH translations database and maintenance system can be used by remote translators to improve the currency and accuracy of their translations and to view and translate new terms as they are added by the MeSH Section throughout the year. The French, Italian, German, and Japanese translators are using the system currently. In January a workshop was held in Vietnam for indexers and translators from several Asian countries. As a result, the Vietnamese have begun a translation, and will be using the system. Others, including the Chinese and Koreans, are considering switching to the system.

Clinical Vocabularies

The MeSH section and its contractors also produce RxNorm, a clinical drug vocabulary that provides standardized names for use in prescribing medications.

RxNorm was designated as a U.S. government-wide target clinical vocabulary standard by the HHS Secretary. It represents the information that is typically known when a drug is prescribed, rather than the specific product and packaging details that are available at the time a medication is purchased or administered. RxNorm provides a mechanism for connecting information from different commercial drug information services. In FY 2005, RxNorm was linked to additional commercial drug terminologies within the UMLS Metathesaurus, and NLM established agreements with the Micromedex, Gold Standard, and Medi-Span drug information providers for regular electronic data feeds to assist in keeping RxNorm up-to-date. LO and OCCS made major improvements in the processing system used to create RxNorm with updates now being issued on a monthly basis and resynchronization of RxNorm files after every Metathesaurus release. The documentation for RxNorm was enhanced and expanded on the NLM Web site.

Through LO's NICHSR, NLM supports the continued development and free distribution of LOINC® (Logical Observation Identifiers Names and Codes) by the Regenstrief Institute. LOINC was designated as a U.S. government-wide target clinical vocabulary standard in 2003, and proposed for adoption as standard under the administrative simplification provision of the Health Insurance Portability and Accountability Act in a draft rule issued in September 2005. NLM continued to manage the contract and pay the annual update fees for the U.S.-wide license for the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT®) in FY 2005.

The existence of authoritative electronic mappings from standard clinical vocabularies to administrative code sets is likely to facilitate automated production of bills and statistical reports as a by-product of the capture of detailed patient data. Mapping efforts must involve both vocabulary producers and intended users, undergo technical review and testing within real clinical systems, and establish effective mechanisms for keeping mappings up-to-date and responding to user feedback. NLM has enlisted cooperation from relevant federal agencies and private organizations and started projects to map LOINC to Current Procedural Terminology (CPT) and SNOMED CT to CPT and to the International Classification of Diseases, 9th edition. The first draft mappings for LOINC to CPT and SNOMED CT to ICD-9-CM will be available in the UMLS Metathesaurus for testing in early 2006.

In September 2004 NLM entered into a three-year contract arrangement with Health Level Seven, Inc. (HL7) to (1) align HL7 message standards with Consolidated Health Informatics standard vocabularies, specifying which subsets of standard vocabularies are valid for particular message segments and replacing

HL7-maintained lists of coded values with subsets of standard vocabularies where feasible (NLM-initiated) and (2) create implementation guide(s) for transmitting an entire Electronic Health Record between disparate systems (on behalf of HHS). As part of this effort, code sets defined for use in version 2 and 3 of HL7 were added to the UMLS Metathesaurus. NLM is coordinating all of its health data standards efforts with related initiatives of the Office of the National Coordinator for Health Information Technology.

UMLS Metathesaurus

The MeSH Section manages the content editing of the UMLS Metathesaurus, using systems developed by the Lister Hill Center. In FY 2005, NLM began to transition responsibility for UMLS Metathesaurus production from the LHC to OCCS and LO. This will be a multi-year process to regularize and streamline the production system and separate it from the Metathesaurus research and development activities. As part of this process, the Bibliographic Services Division assumed responsibility for coordinating production and quality assurance of the UMLS Metathesaurus as well as for UMLS documentation and training. At the close of FY 2005, the Metathesaurus contained more than 1.1 million concepts and 4.7 million concept names from 114 source vocabularies in 17 different languages. (See further information about UMLS activities in the Information Products section of this chapter and the UMLS section of the Lister Hill Center chapter.)

Bibliographic Control

LO creates authoritative indexing and cataloging records for journal articles, books, serial titles, films, pictures, manuscripts, and electronic media, using MeSH to describe their subject content. LO also maintains the NLM Classification, a scheme for arranging physical library collections by subject that is used by health sciences libraries worldwide. NLM's authoritative bibliographic data improves access to the biomedical literature in the Library's own collection, in thousands of other libraries, and in many electronic full-text repositories.

Cataloging

In general, LO adheres to the *Anglo-American Cataloging Rules, 2nd edition*, when creating cataloging and name authority records that reduce the level of cataloging effort required in other libraries. In FY 2005, LO participated in the international review of the draft of proposed cataloging rule changes and played an important role in identifying deficiencies in the proposal which lead to rethinking of these revisions.

In FY 2005, the Technical Services Division's (TSD) Cataloging Section cataloged 21,416 contemporary books, serial titles, non-print items, and cataloging-in-publication galleys. Catalogers began incorporating summary information supplied by participating E-CIP publishers in Cataloging in Publication bibliographic records for medical materials.

The Cataloging Section updated and expanded access to the *NLM Policy on Subject Analysis and Classification*, making it available in HTML and PDF formats. The document is also available on the Web through LC's *Cataloger's Desktop and the Metadata and Cataloging Education Web Clearinghouse*. The 2005 edition of the *NLM Classification* was published in April and included major revisions to the areas of cells and genetics, military medicine, and veterans' health. A well-received new poster of the *NLM Classification* was widely distributed to medical libraries throughout the world. Completion of several special serials cataloging projects eliminated the backlog of new serials acquired by HMD; improved subject access to 100 current and formerly indexed serials; established procedures for handling serials that cease publishing in print and become online-only publications; and facilitated cataloging of microfilmed serials.

The New York Academy of Medicine concluded the project to provide bibliographic records for fugitive literature in health technology assessment in FY 2005, with that organization having contributed a total of 1,219 titles since the project began in 2002.

HMD cataloged 2,668 monographs, 453 linear feet of manuscripts, 2,952 historical audiovisuals, and 2,974 pictures, including a collection of World War II photographs. A project to catalog NLM's important collection of early Japanese items, primarily from the Edo period (1603–1867) was completed. New Profiles in Science Web sites were released for Nobel prize winners Francis Crick (1962), Albert Szent-Gyorgyi (1937), and Salvador Luria (1969). Crick's discovery of the double helix of DNA in 1953 with James Watson is considered the most significant advance in the understanding of biology since Darwin's theory of evolution. Szent-Gyorgyi was recognized for his work in biological oxidation and vitamin C. Luria, a bacteriologist, was honored for discoveries concerning the replication mechanism and the genetic structure of viruses.

Indexing

LO indexed 4,928 biomedical journals for the Medline/PubMed database to assist users in identifying articles on specific biomedical topics. The indexing workload continues to rise, partly due to the selection of new journals for Medline/PubMed, but primarily due to increases in the number of articles published in journals

already being indexed. In FY 2005, a combination of in-house staff, contractors and cooperating U.S. and international institutions indexed 606,000 articles, a 6 percent increase from last year. Previously indexed citations were updated to reflect 67 retractions (that retracted a total of 74 articles), 6,602 corrections, and 30,855 comments found in subsequently published notices or articles. New indexing contracts were awarded on April 1, 2005.

In FY 2005, indexers created 45,312 annotated links between newly indexed MEDLINE citations for articles describing gene function in selected organisms and corresponding gene records in the NCBI Entrez Gene database. This represents a 35 percent increase over the previous year. In June, indexers began to assign new funding support tags to allow monitoring of compliance with the NIH Public Access policy. Research Support, N.I.H., Intramural will be applied when an article explicitly acknowledges, "…supported by the Intramural Research Program of NIH"; Research Support, N.I.H., Extramural will be applied when support from an NIH grant is acknowledged in an article. A third tag, "Wellcome Trust" will be applied when an article explicitly acknowledges support from Wellcome Trust. This information will appear in the MEDLINE Grant Number [GR] field.

Increases in the number of articles and journals indexed for MEDLINE require new approaches to indexing to improve effectiveness and efficiency of the critical and high volume indexing process. In FY 2005 the Index Section continued installing dual monitors for contract indexers to speed indexing from electronic versions of journals. Approximately 20 percent of contract indexers now have dual monitors. A team of staff began assigning subject-appropriate online journals to contractors with dual monitors, including special instructions for navigating through each journal, to reduce the detrimental impact on overall indexing production during the transition period. A new Web-based training program for indexing was used for the first time with the August indexing class and was well received. The class was taught by a combination of the traditional method and the Web-based program. The goal is to eventually use the Web-based program as the only means of training new MEDLINE indexers, which will enable training to occur as staff and contractors are brought on board. LO continues to work with other NLM program areas to identify, test, and implement ways to reduce or eliminate tasks now performed by human indexers. An NLM-wide research and development effort to improve indexing performance and productivity was announced by the Director. The effort is being led by LO.

An NLM Interactive Content Working Group led by LO was formed to recommend ways to identify interactive content that is increasingly appearing in the online journal literature and to describe the existence of this content in MEDLINE citations. This will enable users to easily find articles with accompanying interactive media such as audio, video, manipulable data, animations, etc.

Indexers perform their work after the initial data entry of citations and abstracts is accomplished by one of two means: electronic submission from publishers (the fastest and most economical method), or scanning and optical character recognition (OCR). Work continues to improve the efficiency of data entry. In FY 2005, 76 percent of the citations and abstracts were received electronically from publishers and 24 percent were scanned. The scanning/OCR creation stream was modified to handle all non-English language journals. A total of 385 publishers are now supplying XML-tagged electronic data for 3384 journals.

NLM selects journals for indexing with the advice of the Literature Selection Technical Review Committee (LSTRC) (Appendix 6), an NIH-chartered committee of outside experts. In FY 2005, the Committee reviewed 421 journals and rated 73 highly enough for NLM to begin indexing them immediately. Another 86 ranked sufficiently highly to be indexed if their publishers are able to supply electronic citation and abstract data.

Information Products

NLM produces databases, publications, and Web sites that incorporate authoritative indexing, cataloging, and vocabulary data and link to other sources of biomedical information. LO works with other NLM program areas to produce some of the world's most heavily used biomedical and health information resources.

Databases

LO manages the creation and maintenance of the content of Medline/PubMed, NLM's database of indexed citations; the NLM catalog which is available to the public in two different databases; MedlinePlus and MedlinePlus en español, NLM's primary information resources for patients, families, and the general public; and a number of specialized databases, including several in the fields of health services research, public health, and history of medicine. These databases are richly interlinked with each other and with other important NLM resources, including PubMed Central, other Entrez databases, ClinicalTrials.gov, Genetics Home Reference, as well as the Specialized Information Service's toxicological, environmental health, and AIDS information services.

Use of Medline/PubMed increased to nearly 884 million searches, a 30 percent increase over the

previous year, most directly in PubMed and some via the NLM Gateway. Page views in PubMed totaled 3.2 billion, a 26 percent increase from FY 2004. Google indexing of selected PubMed content has contributed to this growth. Medline/PubMed now includes more than 15.8 million citations. Work progressed on improving subject access to the pre-1966 *Index Medicus* citations which were merged into PubMed in FY 2004 by mapping old MeSH keywords in the records to the 2005 MeSH headings. At the end of the year, mapping was completed for 70 percent of citations.

BSD staff assisted NCBI with the design, development and testing of many enhancements to PubMed, including implementing a Special Queries page and link from PubMed's sidebar which gathers together the various formulated PubMed searches for easy access by users; redesigning the Clinical Queries page; implementing full author and first author searching in Single Citation Matcher, RSS Feed, journal and author autocomplete, Spell Check, replacing Cubby with MyNCBI, and many other features. BSD also worked with LHC on the development and testing of a new version of the Gateway and testing of enhancements to the Gateway's searching of TOXLINE and Hazardous Substances Data Bank (HSDB) and a new search engine.

Use of MedlinePlus and MedlinePlus en español continues to increase significantly. Nearly 75 million unique visitors viewed over 661 million pages. The number of page views increased by more than 25 percent and the number of visitors increased by almost 30 percent. There are more than 70,000 subscribers to the weekly announcements of new content. Both sites remain in the top five of government sites based on the American Customer Satisfaction Index (ACSI). PSD and OCCS continued to expand and improve the basic content and features of the English and Spanish MedlinePlus sites. Thirty new health topic pages were added to MedlinePlus to bring the total to 707; 33 were added to the Spanish site for a total of 658. Nine new interactive tutorials were added in both languages, including highly requested topics on AIDS, C-Section, Vaginal Birth, Cold Sores, and Flashes and Floaters. Within 48 hours of Hurricane Katrina, 27 site records were added to topics on Disaster and Emergency Preparedness, Drinking Water, Food Safety, Food Contamination and Poisoning, Bacterial Infections, and Coping with Disasters. A new video feature of one-hour versions of surgical procedures performed at various hospitals in the U.S., OR-Live, debuted in FY 2005. In addition to archived surgeries, MedlinePlus linked to several live surgeries that allow users to email questions to the surgical team to answer during the procedure. The total number of OR-Live programs available at the end of FY 2005 is 76. Users send many positive comments about OR-live.

With the May 2004 release of the NLM-hosted Go Local system, NLM began serving organizations building Go Local sites with a complete Web-based system that manages records, maps vocabulary, creates HTML, standardizes formats, checks links and provides statistics. In 2005, four new Go Local sites—Indiana, Massachusetts, Alabama, and South Texas—released their collections of health services information to the public using this NLM system. At the annual meeting of the Medical Library Association, NLM sponsored a Go Local sharing session attended by about 60 people. Participants shared lessons learned with current participants and professionals considering embarking on a Go Local project.

NIHSeniorHealth participated in the American Customer Satisfaction Index survey for the first time and received a score which placed it in a tie with NLM's AIDSinfo site for seventh place among government/new sites. Use of the site nearly doubled in FY 2005 with half a million visitors using 6.5 million pages. NLM worked with the National Institute on Aging and other NIH Institutes to bring 13 new health topics to the site. A new feature, Exercise Stories, which includes testimonials of seniors discussing their exercise routines, debuted in FY 2005. PSD staff worked with a usability and design contractor to complete a comprehensive usability evaluation of the site in order to design the next generation of the site to accommodate so many topics. NLM and NIA received an "Industry Innovators" award from the International Council on Active Aging.

Under the direction of NICHSR, NLM continues to expand and enhance its Web-based resources and databases for health services researchers and public health professionals. In FY 2005, NICHSR premiered HSR Information Central, a portal page to resources for health services research; worked with NCBI to expand coverage of health services research and public health information by adding to the Entrez system, as part of the Bookshelf, nearly 40 documents in the HSTAT (Health Services and Technology Assessment); redesigned the interface to HSRProj (Health Services Research Projects; and continued to work through AcademyHealth and the Sheps Center at the University of North Carolina, Chapel Hill to expand HSRProj content. Nearly 900 new records were added and slightly over 2500 records were moved into the HSRProj archive files. Organizations contributing data for the first time in FY 2005 included the National Center on Minority Health and Health Disparities, Illinois Department of Public Health, Michael Reese Health Trust, and Atlantic Philanthropies. The HSRR (Health Services and Sciences Research Resources) database also continued to expand to cover additional datasets, surveys, other research instruments, and software packages used with datasets. A new initiative

at the University of Kentucky, with funding from the Robert Wood Johnson Foundation, is focused on adding to HSRR resources of particular interest for public health systems research.

Machine-Readable Data

NLM leases many of its electronic databases to other organizations to promote the broadest possible use of its authoritative bibliographic, vocabulary, and factual data. There is no charge for any NLM database, but recipients must abide by use conditions which vary depending on the database involved. The commercial companies, International MEDLARS centers, universities, and other organizations that obtain NLM data incorporate them into many different database and software products and use them in a variety of research and development projects.

Demand for Medline/PubMed data in XML format continues to increase. There are currently 330 MEDLINE licensees, an increase of 40 from last year. The new research license program continues to be a success, as 118 non-U.S. organizations use this option. A relatively small number of organizations license NLM catalog records or one or more of the SIS toxicological or environmental health files in XML format. A process was instituted to ensure that deleted catalog records are sent to MARC licensees on a regular basis. Many users execute the online Memorandum of Understanding that permits FTP transfer of the MeSH files in XML, ASCII, or MARC format. At the end of FY 2005, there were 2,933 UMLS Metathesaurus licensees.

Web and Print Publications

NLM's databases and Web sites are its primary method of publication. Demand for the Library's few remaining print publications continues to decline because of increasing electronic access to NLM information and data throughout the world. After 125 years, NLM ceased publication of the monthly *Index Medicus* at the end of calendar 2004. Launched by John Shaw Billings in 1879, *Index Medicus* was an indispensable tool for medical librarians, researchers, and practitioners for decades. The desire to make access to the biomedical literature available in a more timely way was the impetus for NLM's pioneering work in automation in the early 1960s, providing the foundation for the development of MEDLINE in 1971. With the widespread availability of the Internet and the explosion of the biomedical literature, the printed *Index Medicus* had outlived its usefulness. It survives, however, as a searchable subset within Medline/PubMed. The "Black and White" printed MeSH still receives considerable use and continues to be published in print. With the

cessation of *Index Medicus*, the *List of Journals Indexed in Index Medicus* also ceased in 2004 and was replaced in 2005 by the *List of Journals Indexed for MEDLINE*, a comprehensive listing of all journals currently indexed in MEDLINE that includes hundreds of journals that were not in *Index Medicus*.

Following the debut of a complete redesign in FY 2004, NLM's main Web site showed a steady number of 8 million unique visitors viewing 17 percent more pages in FY 2005. Work has focused on the release of a new version of NLM's content management system, anticipated early in FY 2006, including testing, the development of a training manual, and a schedule of hands-on courses for nearly 80 contributors. PSD and OCCS have also been reviewing candidates to replace the main site, archive, and MedlinePlus search engines.

Publications available from the main Web site include recurring newsletters and bulletins, fact sheets, technical reports, and documentation for NLM databases. Following a total revision in FY 2004, TSD published the first fully searchable 4th edition of the *Collection Development Manual of the National Library of Medicine* on the Web. BSD's MEDLARS Management Section edits the NLM *Technical Bulletin* which provides timely, detailed information about changes and additions to NLM's databases and related policies, primarily for librarians and other information professionals. Published since 1969, it also serves as the historical record of the evolution of NLM's online systems and databases. Several improvements to the functionality of the *Technical Bulletin* became available in FY 2005 including the ability to print an entire issue, the ability to search article text back to December 1997 and the index back to 1979, and a feature available on the current issue which shows the articles that are most read by users.

Direct User Services

In addition to building databases and producing other heavily used electronic information products, LO provides document delivery and reference and customer services to remote users as a national and international backup to services available from other health sciences libraries and information suppliers. LO also serves onsite clientele in the NLM reading rooms.

Document Delivery

LO retrieves materials requested by onsite patrons from NLM's closed stacks and also provides interlibrary loan as a backup to document delivery services available from other libraries and information suppliers. In FY 2005, PSD's Collection Access Section processed 580,072 requests for contemporary documents. HMD handled 13,384 requests for historical documents, rare

books, manuscripts, pictures, and historical audiovisuals.

The number of onsite requests in the NLM Main Reading Room continues to decline due to in part to increased security measures, but primarily due to online access to more journals in the Reading Room and free access to journals in PubMed Central. Onsite users requested 238,833 contemporary items from the stacks, a 12 percent decline from last year. On the other hand, users of the HMD Reading Room requested 11,940 items from historical and special collections, a 28 percent increase from last year attributed to increases in requests for prints and photographs and historical audiovisuals. A new four-year contract for photocopy, printing and patron registration was implemented which includes fully networked black and white printing and net-workable digital photocopiers than can also serve as copier/printers if the needs of the Library change. A study of patron usage patterns resulted in changes in the hours the Reading Room is open. Thursday evening hours were eliminated and Saturday hours were extended from 12:30 to 2:30 pm. Patron limits on daily stack requests were raised from 16 to 50, which has had little impact on stack requests due to the availability of many electronic journals.

In FY 2005, PSD's Collection Access Section processed a total of 341,239 interlibrary loan requests, a 5 percent decline from the previous year. A 2 percent rise in the fill rate, to a new high of 80 percent was due in part to continued collaboration with the Index Section to make more indexed titles available for document delivery. NLM now delivers 96 percent of interlibrary loan requests electronically.

A total of 3,234 libraries use DOCLINE, NLM's interlibrary loan request and routing system. NLM released DOCLINE 2.5 and Loansome Doc 2.5, unveiling a totally redesigned appearance for Loansome Doc, many new features, and an easier to use interface. Highlights of the release include password reset feature, display of library holdings during the order process, a warning message of possible duplicate order, email confirmation of registration, ability for patrons to email their library from within Loansome Doc, and the ability for libraries to edit their patron records. SSL (secure socket layer) was also added to Loansome Doc to provide security of patron information as it is passed across the Internet. DOCLINE users entered 2.48 million requests in FY 2005, a 9.5 percent decrease from the previous year; 91 percent of the requests were filled. Although the absolute number of interlibrary loan requests received by NLM declined in FY 2005, the Library's share of all DOCLINE requests continues to increase by about half a percent each year—to 13.8 percent in FY 2005. Individuals submitted 677,976 requests to DOCLINE libraries via the Loansome Doc feature in Medline/PubMed and the NLM Gateway, a

16 percent decline from the previous year. Document request traffic continues to decline in all Regions of the NN/LM due to expanded availability of electronic full text journals.

NCBI and staff at the Regional Medical Libraries continued to promote the use of PubMed's LinkOut for Libraries and "Outside Tool" as a means for libraries to customize PubMed to display their electronic and print holdings to their users. The number of libraries participating in LinkOut increased to 1,353, 24 percent more than in the previous year.

DOCLINE requests are routed to libraries automatically based on holdings data in its serial holdings database. At the end of FY 2005, the holdings database contained 1.4 million holdings statements for 53,857 serial titles held by 3,041 libraries. In FY 2005 the OCLC to DOCLINE serials holdings batch load program and the ability to provide holdings in MARC 21 format were implemented.

NLM and the Regional Medical Libraries continued to encourage network libraries to use the Electronic Funds Transfer System operated for the NN/LM by the University of Connecticut to reduce administrative costs associated with ILL billing. During FY 2005, Electronic Funds Transfer participation increased 11 percent to 1,068 libraries. Participants receive either a single net consolidated bill or a net consolidated payment each month.

In the aftermath of Hurricane Katrina, NLM made the decision not to charge for interlibrary loans provided to libraries affected by the hurricane during July-September 2005. The RMLs identified 65 libraries that either had significant losses to their collections or staff or were providing library service to users of libraries which remained closed. A total of 502 requests were not billed to affected institutions. NLM and the RMLs will evaluate the situation to see if continued assistance is needed for the October–December 2005 billing period. Routing of requests to affected libraries was also disabled.

Reference and Customer Service

LO provides reference and research assistance to onsite and remote users as a backup to services available from other health sciences libraries. LO also has primary responsibility for responding to inquiries from those seeking information about NLM's products or services or assistance in using these services. PSD's Reference and Web Services Section handles all initial inquiries with contract assistance and many of those requiring second-level attention. Staff throughout LO and NLM assist with second-level service when their specialized expertise is required.

In FY 2005, the Reference and Web Services Section handled a total of 96,000 customer inquiries

from onsite and remote patrons, an 11 percent decrease from the previous year. Onsite requests declined 27 percent, but remote requests increased 8 percent. Cosmo, NLM's virtual customer service representative responded to 4,600 questions, 24 percent more than in FY 2004; 90 percent of in-scope requests were answered correctly. Questions that Cosmo is unable to answer are transferred to Reference staff at the user's request.

Outreach

LO manages or contributes to many programs designed to increase awareness and use of NLM's collections, programs, and services by librarians and other health information professionals, historians of medicine and science, researchers, educators, health professionals, and the general public. LO coordinates the National Network of Libraries of Medicine which works to equalize access to health information services and information technology for health professionals and the public throughout the United States; serves as the secretariat for the Partners in Information Access for the Public Health Workforce; participates in NLM-wide efforts to develop and evaluate outreach programs designed to improve information access for underserved minorities and the general public; produces major exhibitions and other special programs in the history of medicine; and conducts a range of training programs for health sciences librarians and other professionals. LO staff members give presentations, demonstrations, and classes at professional meetings and publish articles to highlight NLM programs and services.

National Network of Libraries of Medicine

The NN/LM works to provide timely, convenient access to biomedical and health information resources for U.S. health professionals, researchers, educators, and the general public, irrespective of their geographical location. It is the core component of NLM's outreach program and its efforts to reduce health disparities and improve health information literacy. The network includes 5,794 full and affiliate members. The full members are libraries with health sciences collections, primarily in hospitals and academic medical centers. Affiliate members include some small hospitals, public libraries, and community organizations that provide health information services, but have little or no collection of health sciences literature. LO's NN/LM Office oversees the network programs that are administered by eight Regional Medical Libraries under contract to NLM. As mentioned previously, the 5-year NN/LM contracts are currently being recompeted. (See Appendix 1 for a list of the RMLs.)

In addition to the basic NN/LM contracts, NLM funds subcontracts for four national centers that serve the entire network. The activities of one of these centers, the National Online Training Center and Clearinghouse at the New York Academy of Medicine, are described elsewhere in this chapter. The Outreach Evaluation Resource Center at the University of Washington provides training and consulting services throughout the NN/LM and assists NLM, the RMLs, and other network members in designing methods for measuring the effectiveness of overall network programs and individual outreach projects. In FY 2005, the Center focused on completing the evaluation of NN/LM outreach efforts to public libraries and the public health workforce.

The National Outreach Mapping Center at Indiana University in Indianapolis assists NLM in displaying visually the geographic distribution and impact of NN/LM programs and services. In FY 2005 the Center worked on expanding the mapping and reporting application to depict locations of project lead institutions. Work continued on collecting uniform outreach activity data from all Regions. The Web-Services Technology Operations Center (Web-STOC) provides ongoing technical management of the NN/LM Web sites and also investigates, recommends, and directs the implementation of additional Web technology for teleconferencing, Web broadcasting, distance education, online surveys, etc. A major activity for the Web-STOC in FY 2005 was assisting with the development of a Request for Proposals and the evaluation and selection of a contractor to evaluate and redesign the NN/LM main Web site and the regional sites of six RMLs. The American Institutes for Research received the contract, which was awarded through the RML at the University of Washington, home of the Web-STOC.

The RMLs and other network members conduct many special projects to reach underserved health professionals and to improve the public's access to high quality health information. Most of these projects involve partnerships between health sciences libraries and other organizations, including public libraries, public health departments, professional associations, schools, churches, and other community based organizations. Some projects are identified by individual RMLs through regional solicitations or ongoing interactions with regional institutions; others are identified by periodic national solicitations for outreach proposals issued simultaneously in all NN/LM regions. In all, the NN/LM issued 153 subcontracts for outreach projects in FY 2005. Many of these projects focused on improving access to information for consumers and the public health workforce.

With the assistance of other NN/LM members, the RMLs conduct most of the exhibits and

demonstrations of NLM products and services at health professional, consumer health, and general library association meetings around the country. LO organizes the exhibits at the Medical Library Association annual meeting, the American Library Association annual meeting, some of the health professional and library meetings in the Washington, DC area, and some distant meetings focused on health services research, public health, and history of medicine. In FY 2005, NLM and NN/LM services were exhibited at 235 national, regional, and state meetings across the U.S. These exhibits highlight all NLM services relevant to attendees.

As a result of input from network members at site visits to the eight RMLs in 2002–2003, NLM and the RMLs established an NN/LM Hospital Internet Access Task Force in FY 2003 to identify barriers to access to the Internet in hospitals, best practices for achieving the twin goals of easy access to the Internet and appropriate security for hospital patient data, and actions the NN/LM and NLM might take to assist hospital libraries in overcoming barriers. The Task Force continued to work with the Hospital Libraries Section of MLA to promote a new "Information Technology Forum" and in FY 2005 distributed its final report to the NN/LM via regional listservs.

Also as a result of input from the site visits, NLM and the RMLs established an E-licensing Working Group to identify state and local group licensing resources available to network members, model licensing language, best practices for negotiating licenses, and methods for disseminating the information to network members. The group continued its work in FY 2005 in collaboration with MLA, identifying several resources that are standard licensing models which may be useful to hospital libraries in their negotiations and recommending that RMLs assist smaller libraries with developing group licenses. A final report with additional recommendations is expected in FY2006.

Partners in Information Access for the Public Health Workforce

The NN/LM is a key member of the Partners in Information Access for the Public Health Workforce, a 12-member public-private agency collaboration initiated by NLM, the Centers for Disease Control and Prevention, and the NN/LM in 1997 to help the public health workforce make effective use of electronic information sources and to equip health sciences librarians to provide better service to the public health community. The NICHSR coordinates the Partners for NLM; staff members from the National Network Office, SIS, and the Office of the Associate Director for Library Operations serve on the Steering Committee, as do representatives from several RMLs.

The Partners Web site (PHPartners.org), managed by NLM with assistance from the Massachusetts RML, provides unified access to public health information resources produced by all members of the Partnership, as well as other reputable organizations. In FY 2005, the Web site was expanded significantly, with more than 200 new links added, and a weekly email notification service was started. Work is nearly complete on a series of topic pages to be added to the site. One of the most popular resources already on the site is the Healthy People 2010 Information Access Project. For every focus area of Healthy People 2010, the Project resource includes four or more objective-specific evidence-based PubMed search strategies and links to MedlinePlus topics. In addition, the Project site was evaluated this year by the Public Health Foundation and numerous improvements were made in response to the recommendations which resulted.

The Partnership also devoted considerable effort to the development of training resources. *Public Health Information and Data: A Training Manual*, developed collaboratively by several agencies in FY 2004, was converted in FY 2005 by University of Michigan, NICHSR, and NNO staff into a Web-based tutorial and made available at http://phpartners.org/tutorial/index.html.

In addition, in FY 2005 NICHSR awarded six purchase orders under the Partnership to the: Association of Schools of Public Health for *Update and Improved Usability of Distance Learning Information Contained in ASPH Website for Public Health Workforce* ; Association of State and Territorial Health Officials for *Strategies for Enumerating the Public Health Workforce* (http://www.astho.org/pubs/WorkforceEnumerationRep ort.pdf) Delta Omega for *Public Health Reports: Historical Collection 1878–2005*; Medical Library Association for a March 2005 satellite broadcast on *Partnering for Public Health: Information, Librarians, and the Public Health Workforce* (http://www.mlanet.org/education/telecon/publichealth); National Association of County and City Health Official for *Analysis of Local Public Health Agency Information Technology Infrastructure*; and Society for Public Health Education for *CDCynergy: Diffusion of an Informatics Innovation to Strengthen the Public Health Workforce and Improve Public Health.*

Special NLM Outreach Initiatives

LO participates in many NLM-wide efforts to expand outreach and services to the general public and to address racial and ethnic disparities and participates actively in the Library's Committee on Outreach, Consumer Health, and Health Disparities. In December 2004, NLM sponsored a Community-Based Outreach

Symposium, planned by staff from the NN/LM Office, the Office of Health Information Programs Development, Specialized Information Services, and the RMLs. This Symposium is described in the Office of Health Information Programs Development (OHIPD) chapter, but the insights gained in the Symposium helped LO to shape the outreach portions in the RFP for the NN/LM contract for 2006–2011.

HMD worked with OHIPD to establish an arrangement with MHA Enterprises, an American Indian business, to provide scanning and XML data capture services for the Library's archives. In collaboration with OHIPD and a contractor, they conducted training on the reservation near Parshall, ND. Among the documents scanned were the minutes of NLM Board of Regents meetings, NLM Annual Reports, and back volumes of the NLM Technical Bulletin, which were made publicly available on HMD's NLM Archives Web site.

LO also worked with other NLM components, the American College of Physicians, and some NN/LM members to continue the Information Rx project to test the use of "information prescriptions" for MedlinePlus in physicians' offices throughout the U.S. The Office of the Associate Director, LO, the NNO, and BSD continued to work with the Public Library Association and the American Library Association to improve public library awareness of MedlinePlus and MedlinePlus en español. The Office of the Associate Director also serves on an ALA Advisory Committee for "Be Well Informed@Your Library" which funded 10 public library systems to conduct seminars on health education.

LO staff members continue to be involved in NLM's partnership with the SCIMATCEH Academy at Wilson High School in the District of Columbia. In FY 2005, LO provided summer employment and training opportunities for several students.

Historical Exhibitions and Programs

HMD directs the development and installation of major historical exhibitions in the NLM rotunda, with assistance from LHC and the Office of the Director. Designed to appeal to the interested public as well as the specialist, these exhibitions highlight the Library's historical resources and are an important part of NLM's outreach program. The current exhibition, *Changing the Face of Medicine: Celebrating America's Women Physicians,* debuted on October 14, 2003 with a gala opening program featuring remarks by Dr. Elias Zerhouni and other dignitaries.

This widely reviewed and highly praised exhibition features more than 300 women physicians, living and dead, selected with advice from an advisory committee of eminent physicians (both women and men), chaired by Tenley Albright, M.D., former chair of the NLM Board of Regents. Girls who might be interested in pursuing an M.D. degree are one of the principal audiences for the exhibition, which illustrates the wide range of careers open to women physicians and shows that women from all segments of U.S. society have excelled in the field. The exhibition has a Web site that provides information about the women physicians in the exhibition and educational and professional resources for those considering a career in medicine. The "Share Your Story" section of the Web site encourages people to provide information about outstanding women physicians they have encountered, whether family members, mentors, or their own doctors. To date, approximately 17,000 visitors have seen the exhibition at NLM and the accompanying Web site has received 1.9 million page views. In March 2005 HMD hosted a two-day symposium in conjunction with the exhibition on "Women Physicians, Women's Politics, Women's Health." Papers were presented on subjects such as the history of the women's health movement, images of women physicians, and race/class/gender issues in medicine.

NIH gave its Health Plain Language award to the exhibition text, brochure, Web site and calendar. NLM and the American Library Association completed design and production for a traveling version of *Changing the Face*, funded by the NIH Office of Research on Women's Health and the Library. Sixty libraries will host the traveling exhibition over the next five years.

While *Changing the Face* was on display, work continued on the next major exhibition, *Visible Proofs: Forensic Views of the Body*, which will open in February 2006. *Visible Proofs* explores the developments in scientific methods that translate views of bodies and body parts into visible proofs. It tells stories of the people, sciences, and technologies that make visible the cause and manner behind a death. Visitors to the exhibition will observe, analyze, and decipher different forensic views of the body. Visitors will examine important historical and contemporary cases and forensic techniques through the use of objects, graphics, and multimedia presentations. They will also encounter forensic experts whose contributions and discoveries have changed the field of forensic medicine.

Previous NLM exhibitions live on through heavily used Web sites, printed catalogs, DVDs, or touring traveling versions. Excluding *Changing the Face*, exhibition Web sites received more than 8.6 million page hits in FY 2005. The traveling version of *Frankenstein: Penetrating the Secrets of Nature* continued its multi-year tour of public, academic, and health sciences libraries across the U.S. under the auspices of the American Library Association and continued to draw large crowds and favorable reviews.

In addition to the major exhibitions in the rotunda, HMD installs "mini-exhibits" generally in cases near the entrance to the HMD Reading Room. An exhibit on *Strange Hells within the Minds War Made: War and Trauma in the 20th Century* was on display November 9, 2004–May 31, 2005. The exhibit focused on World War I shell shock, delineating medical and literary responses to the war and included a remarkable collection of artifacts, books, and images. It then moved to the National Museum of Health and Medicine at Walter Reed Army Medical Center where it was on display until August 13. On June 13, *50 Years of the American College of Nurse-Midwives* went on display. It was toured by attendees of the College's annual meeting in Washington. A mini-exhibit on the life and career of long-time NLM staff member Stanley Jablonski opened on August 29. Stanley worked for NLM from 1949 to 1976, most of that time in the Index Section where he rose to Head of the section. He continued to work at NLM, as a scholar, for nearly 30 years after his retirement. The exhibit highlights his career as an indexer and author and his passion for sailing.

In November 2004, HMD sponsored a two-day workshop on Cancer in the Twentieth Century in collaboration with The Centre for the History of Science, Technology and Medicine (University of Manchester) and the Society for the Social History of Medicine. The purpose of the workshop was to prepare teaching resources for the history of cancer suitable for use in the university curriculum. This is part of an ongoing program at NLM designed to encourage the incorporation of HMD research materials in university classrooms and includes the development of Web syllabi and course materials. Other historical programs include a monthly series of seminars by historical scholars and several special historical lectures organized by HMD in conjunction with the Diversity Council and the EEO Office. HMD also hosted a number of visiting historical scholars,

HMD staff members continued to present historical papers at professional meetings and to publish the results of their scholarship in books, chapters, articles, and reviews. HMD continued to play a lead role in preparing the recurring features "Voices from the Past" and "Images of Health" for the *American Journal of Public Health*, which often features materials from the NLM collection.

Training and Recruitment of Health Sciences Librarians

LO develops online training programs in the use of Medline/PubMed and other databases for health sciences librarians and other search intermediaries; oversees the activities of the National Online Training Center and Clearinghouse (NTCC) at the New York

Academy of Medicine; directs the NLM Associate Fellowship program for post-Masters librarians; and develops and presents continuing education programs for librarians and others in health services research, public health, the UMLS resources, and other topics. LO also collaborates with the Medical Library Association, the Association of Academic Health Sciences Libraries, and the Association of Research Libraries to increase the diversity of those entering the profession, to provide leadership development opportunities, to promote multi-institution evaluation of library services, and to encourage specialist roles for health sciences librarians.

In FY 2005, the MEDLARS Management Section (MMS) and the NTCC trained 958 students in 69 classes covering PubMed, the Gateway, ClinicalTrials.gov, TOXNET, and the UMLS. A new version of the PubMed Tutorial which included the My NCBI feature was implemented and new Quick Tours of MyNCBI and the NLM Catalog were introduced. The PubMed Tutorial was used nearly 43,000 times each month, representing a 7 percent increase in use over last year.

The UMLS courses are one of a number of NLM training courses useful in preparing librarians for new and expanded roles. The courses were updated in FY 2005 to cover enhancements to the UMLS, how licensees are currently using the UMLS in health care systems and research, and potential future uses relating to the promotion and use of standard clinical vocabularies in electronic health records. In addition to teaching the UMLS at the MLA and American Medical Informatics Association (AMIA) Annual Meetings, MMS staff taught the UMLS Basics course at the University of Pittsburgh Library School, New York Academy of Medicine, and at NLM several times throughout the year.

LO and the NTCC assist NCBI in arranging network locations, scheduling, and publicizing the Introduction to Molecular Biology Information Resources class, which helps to prepare library-based bioinformatics specialists. NCBI also offers an advanced workshop for Bioinformatics Information Specialists at NLM. Both courses were developed and are taught by librarians who serve as bioinformatics specialists in universities and at NLM. NICHSR continues to add to its suite of courses on health services research, public health, and health policy.

The NLM Associate Fellowship program had 13 participants in FY 2005: six 2nd year Associates at sites across the country and seven 1st year Fellows who completed their year at NLM in August 2005. Four of the latter also chose to participate in the optional 2nd year of the program at the following sites: the University of Rochester, the University of Maryland, Virginia Commonwealth University, and the University

of Washington. Six new Fellows began a year at NLM in September, including one International Fellow from the library of the Ministry of Health in Mozambique. NLM works with several organizations on librarian recruitment and leadership development initiatives. Individuals from minority groups continue to be underrepresented in the library profession and a high percentage of current library leaders will retire within the next 5 to 10 years. LO has provided support for scholarships for minority students available through the American Library Association, Medical Library Association, and the Association for Research Libraries. LO also supports the NLM/AAHSL Leadership Development Program, which provides leadership training, mentorship, and site visits to the mentor's institution for an annual cohort of five mid-career health sciences librarians. Association of Academic Health Sciences Libraries (AAHSL) contracts with ARL for the leadership training portion of the program.

Table 1

Growth of Collections

Collection	Previous Total (9/30/04)	Added FY 2005	New Total (9/30/05)
Book Materials			
Monographs:			
Before 1500	591	0	591
1501-1600	5,963	13	5,976
1601-1700	10,234	11	10,245
1701-1800	24,655	29	24,684
1801-1870	41,460	122	41,582
Americana	2,341	0	2,341
1871-Present	742,296	15,854	758,150
Theses (historical)	281,794	6,297	288,091
Pamphlets	172,021	0	172,021
Bound serial volumes	1,288,842	18,992	1,307,834
Volumes withdrawn	(87,612)	(11,789)	(99,401)
Total volumes	2,482,585	29,529	2,512,114
Nonbook Materials			
Microforms:			
Reels of microfilm	142,067	3,727	145,794
Number of microfiche	450,403	2,306	452,709
Total microforms	592,470	6,033	598,503
Audiovisuals	74,703	2,641	77,344
Computer software	2,375	94	2,469
Pictures	60,432	2,823	63,255
Manuscripts	4,739,682	572,600	5,312,282*
Total nonbook	5,469,662	584,191	6,053,853
Total book & nonbook	**7,952,247**	**613,720**	**8,565,967**

* Equivalent to 3,035 linear feet.

Table 2

Acquisition Statistics

Acquisitions	FY 2003	FY 2004	FY 2005
Serial titles received	20,476	20,769	20,989
Publications processed:			
Serial pieces	134,579	132,192	132,347
Other	24,523	24,323	24,659
Total	159,102	156,515	157,006
Obligations for:			
Publications	$6,217,417	$6,942,747	$8,255,443
(For rare books)	($297,894)	($300,831)	($324,398)

Table 3

Cataloging Statistics

	FY 2003	FY 2004	FY 2005
Completed Cataloging	19,927	21,238	21,416

Table 4

Bibliographic Services

Services	FY 2003	FY 2004	FY 2005
Citations published in MEDLINE	526,338	571,069	606,597
For *Index Medicus*	492,911	537,469	*ceased publication*
Journals indexed for MEDLINE	4,697	4,839	4,928
Journals indexed for *Index Medicus*	3,994	4,189	*ceased publication*
Articles in PubMed Central	109,910	347,680	432,992

Table 5

Consumer Web Services

Services	FY 2003	FY 2004	FY 2005
NLM Web Home Page			
Page Views	37,166,023	48,335,875	56,332,002
Unique Visitors	4,792,482	7,934,966	7,996,621
MedlinePlus			
Page Views	214,127,932	498,702,940	661,000,000
Unique Visitors	16,356,444	51,724,958	74,600,000
ClinicalTrials.gov			
Page Views	26,010,359	33,651,851	61,303,796
Unique Visitors	2,387,487	3,190,813	3,499,091
Genetics Home Reference			
Page Views	***	6,298,501*	13,633,547
Unique Visitors (daily average)	***	1,631*	3,816
Household Products Database			
Page Views	***	7,096,664	10,547,690
Unique Visitors		1,364,649	1,926,321
Tox Town			
Page Views	***	1,732,336	3,232,466
Unique Visitors		110,660*	187,544

*Corrected figure

Table 6

Circulation Statistics

Activity	FY 2003	FY 2004	FY 2005
Requests Received	653,916	631,806	580,072
Interlibrary Loan	363,352	359,577	341,239
Onsite	290,564	272,229	238,833
Requests Filled:	511,032	510,751	475,623
Interlibrary Loan	268,714	281,543	273,870
Onsite	242,318	229,208	201,753

Table 7

Online Searches—PubMed and NLM Gateway

	FY 2003	FY 2004	FY 2005
Total online searches	504,000,000	678,000,000	754,372,598

Table 8

Reference and Customer Services

Activity	FY 2003	FY 2004	FY 2005
Offsite requests	64,010	71,290	73,493
Onsite requests	41,774	36,649	22,298
Total	105,784	107,939	95,791

Table 9

Preservation Activities

Activity	FY 2003	FY 2004	FY 2005
Volumes bound	15,646	18,311	18,417
Volumes microfilmed	2,795	2,603	1,564
Volumes repaired onsite	1,285	1,652	2,095
Audiovisuals preserved	500	795	936
Historical volumes conserved	111	197	140

Table 10

History of Medicine Activities

Activity	FY 2003	FY 2004	FY 2005
Acquisitions:			
Books	314	498	1,070
Modern manuscripts	498,750	5,516,000*	1,872,500*
Prints and photographs	1,000	1,591	11,252
Historical audiovisuals	97	757	176
Processing:			
Books cataloged	215	13,621	2,668
Modern manuscripts cataloged	203,000	740,250*	792,750*
Pictures cataloged	1,048	2,758	2,974
Citations indexed	856	876**	452
Public Services:			
Reference questions answered	14,693	18,701	20,655
Onsite requests filled	16,163	8,618	12,738

*Estimated figure based on 1750 manuscripts per linear foot.
**Corrected figure

SPECIALIZED INFORMATION SERVICES

Jack Snyder, M.D., J.D., Ph.D.
Associate Director

The Toxicology and Environmental Health Information Program (TEHIP), known originally as the Toxicology Information Program, was established 38 years ago within NLM's Division of Specialized Information Services (SIS). Over the years TEHIP has provided for the increasing need for toxicological and environmental health information by taking advantage of new computer and communication technologies to provide more rapid and effective access to a wider audience. We continue to explore ways to point and link users to relevant sources of toxicological and environmental health information, governmental and nongovernmental, wherever these sources may reside.

Resources include chemical and environmental health databases and Web-based information collections. Development of HIV/AIDS information resources is a focus of the Division and now includes several collaborative efforts in information development and deployment, including an emphasis on the information needs of other special populations. Our outreach program has continued to evolve and reach out to under-served communities through implementation of innovative information access-enabling approaches and dissemination of NLM's resources.

The SIS Web server provides a central point of access for the varied programs, activities, and services of the Division. Through this server (http://sis.nlm.nih.gov), users can access interactive retrieval services in toxicology and environmental health, HIV/AIDS information, and special population health information; find program descriptions and documentation; and be connected to outside related sources. Continuous refinements and additions to our Web-based systems are made to allow easy access to the wide range of information collected by this Division. Our usage has continued to increase over the past year with access to all toxicology and HIV/AIDS data free over the Internet.

In FY 2005 SIS continued to balance efforts to enhance and re-engineer existing information resources with efforts to provide new services in emerging areas. We further developed various prototypes that rely on geographic information systems, innovative access and interfaces for consumers, and graphical display of data from information sources. Highlights for 2005 are in the following sections.

Toxicology and Environmental Health Resources

The TOXNET (TOXicology Data NETwork) is a cluster of databases covering toxicology, hazardous chemicals, environmental health and related areas. These databases continue to be a highly used resource, and in FY 2005 customer surveys 87 percent of the responders reported that they would "return to this site" and "recommend it to others." Enhancements to TOXNET were based on user feedback/requests and routine upgrades/additions of data and capabilities. Databases in TOXNET include:

- **HSDB** (Hazardous Substances Data Base), a file of toxicology information on over 4,900 potentially hazardous chemicals.
- **IRIS** (Integrated Risk Information System), a database from the Environmental Protection Agency (EPA) containing carcinogenic and non-carcinogenic health risk information on over 500 chemicals.
- **ITER** (International Toxicity Estimates for Risk), a database containing data in support of human health risk assessments. It is compiled by Toxicology Excellence for Risk Assessment and contains over 600 chemical records.
- **CCRIS** (Chemical Carcinogenesis Research Information System), a scientifically evaluated and fully referenced data bank, developed and maintained by the National Cancer Institute (NCI) with over 8,900 chemical records with carcinogenicity, mutagenicity, tumor promotion, and tumor inhibition test results.
- **GENE-TOX** (Genetic Toxicology), a toxicology database created by the EPA containing genetic toxicology test results on over 3,000 chemicals.
- **TOXLINE**, a bibliographic database providing comprehensive coverage of the biochemical, pharmacological, physiological, and toxicological effects of drugs and other chemicals from 1965 to the present. TOXLINE contains over 3 million citations, almost all with abstracts and/or index terms and CAS Registry Numbers.
- **DART/ETIC** (Development and Reproductive Toxicology/Environmental Teratology Information Center), a bibliographic database covering literature on reproductive and developmental toxicology.
- **Toxics Release Inventory** (TRI), a series of databases that describe the releases of toxic chemicals into the environment annually for the years 1987–2003.

- **ChemIDplus**, a database providing access to structure and nomenclature authority used for the identification of chemical substances cited in NLM databases. ChemIDplus contains over 368,000 chemical records, of which over 247,000 include chemical structures.
- **Household Products Database**, which provides information on the potential health effects of chemicals contained in more than 6,000 common household products.
- **Haz-Map**, an occupational toxicology database designed primarily for health and safety professionals, but also for consumers seeking information about the health effects of exposure to chemicals and biologics at work. It links jobs and hazardous tasks with occupational diseases and their symptoms.
- **ALTBIB**, a bibliographic database on alternatives to the use of live vertebrates in biomedical research and testing.

The redesign of the **ChemIDplus** application's user interface has significantly enhanced the functionality of this chemical database. The new interface has a simpler search feature for novice users and an advanced search capability for experienced users. A chemical spellchecker has been integrated into the search engine to help non-chemist users. New searchable features in the advanced version include search and display by toxicity indicators such as Median Lethal Dose (LD50), by Physical/Chemical Properties such as LogP, and by Molecular Weight. ChemIDplus records are one of several sources that contribute names, structures, and toxicology information to the PubChem databases that are part of NIH's Molecular Libraries Roadmap Initiative.

WISER (Wireless Information System for Emergency Responders) is a tool developed for use by emergency responders during hazardous materials incidents, as well as during training sessions/exercises in preparation for such events. WISER is now available for downloading onto PDAs (Palm OS and Pocket PC) and Windows laptops/desktops. Through workshops, email feedback, and an online bulletin board focus group, user evaluation of WISER has been very positive and included suggestions for new features and data for future versions. There have been over 25,000 downloads of WISER to PDAs and Windows desktops. Tox Town developed and deployed two new graphic scenes, Tox Farm and a U.S. Mexico Border scene. Unique environmental hazards of these sites are highlighted, along with a common set of potential hazards common to all the scenes. For example, Tox Farm includes crop fields, agricultural runoff, animal waste, farm animals, barn/silos and meat processing. U.S.–Mexico Border scene includes locations such as maquiladora, crop field, colonia, trash burning, and storm water/sewage.

TOXMAP is a Geographic Information System (GIS) system that uses maps of the U.S. to help users visually view data about chemicals released into the environment and easily connect to related environmental health information. Improvements in FY 2005 included a chemical spellchecker, an improved Frequently Asked Questions module, and addition of graphs to present release facility amounts and release trends since 1987.

TOXNET, DIRLINE and Health Hotlines were configured and reformatted for optimum use with PDA devices.

Within a few days after the disaster of hurricane Katrina, SIS compiled and posted a Web site on "Hurricanes: Links to health information including toxicology and environmental health." This Hurricane site provides information on hurricane recovery, cleanup, and handling hazardous substances. It was featured by Science.gov and within days, more than 1,400 other sites linked to it.

Other new Enviro-Health Link pages this year include EdCoTox, links to resources on education, careers, and outreach in toxicology and environmental health; Indoor Air Pollution with information links related to the many sources of indoor air pollution including cleaning and maintenance products, building materials, tobacco smoke, mold, and poor ventilation; Outdoor Air Pollution with relevant information links; and Toxicogenomics, a new subdiscipline that combines the emerging technologies of genomics and bioinformatics to identify and characterize mechanisms of action of known and suspected hazardous substances.

AIDS Information Services

SIS is the project manager for the multi-agency service titled "AIDSinfo." This service provides access to AIDS-related clinical trials (through Clinicaltrials.gov), federally approved treatment guidelines, and prevention and research information. The contract for this service also provides support services for ClinicalTrials.gov and Genetics Home Reference.

The HHS AIDSinfo celebrated World AIDS Day with the unveiling of a new AIDSinfo Web site design. The new design offers a more user-friendly layout and consistent navigation features. It also includes a new service, Live Help, which allows Web site visitors confidential communication with staff in real time to gain assistance with navigating for HIV/AIDS information resources.

NLM offers competitive awards to community-based organizations and libraries to design and conduct projects that will improve access to HIV/AIDS related health information for patients, the affected community,

and their caregivers. Projects must involve one or more of the following information access categories: information retrieval, skills development, state-of-the-art resources, resource development, and document access. In FY 2005, NLM made 13 Standard Awards for up to $50,000 and five Express Awards for up to $10,000.

Evaluation Activities

Many SIS Web pages were redesigned in FY 2005 with a focus on improved navigation and information organization. They were also upgraded to align with the new NLM Web page redesign. Eight SIS Web products were assessed via online surveys, focus groups and or online bulletin forums. Feedback was used as input for enhancement and direction for new capability. Products evaluated were WISER, TOXMAP, Tox Town, TOXNET, ChemIDplus, and the Special Population web sites (Asian American, Arctic Health, and American Indian).

The American Customer Satisfaction Index (ACSI) continues to be used to evaluate TOXNET and AIDSinfo. The 2005 score for TOXNET is 76, which ranks it seventh out of 30 "Information/News Websites." AIDSinfo, which has gone up four points in FY 2005, has a score of 80 and is ranked fourth. This represents the largest quarter-over-quarter increase reported for all government sites, and likely reflects the response to the redesigned AIDSinfo web pages.

Several of the SIS web applications were approved and received certification by the Health On the Net Foundation as adhering to their international code principles. Approved applications include Tox Town, ToxMap, Toxicology and Environmental Health World Library, and AIDSinfo. Other SIS web applications will be submitted for Health on the Net certification in 2006.

Outreach Initiatives

SIS outreach programs reach health professionals, public health workers, and the general public especially about health issues that disproportionately impact minorities, such as environmental exposures and AIDS. The following highlight some of the advances in FY 2005.

The United Negro College fund Special Programs Corporation awarded four Historically Black Colleges and Universities (HCBU) small grants to develop and implement projects that help to increase the awareness and utilization of NLM resources on campuses and in their communities. This program is now in its third year, and the evaluation reports from earlier grants are providing evidence of successful implementations.

The Adopt-a-School program with Woodrow Wilson Senior High School (Washington, D.C.) encourages students to take an active interest in consumer health and promotes interest in science. Projects this year included online training about NLM databases, summer internships for students, donation of technological books and periodicals, NLM tours, and guest lectures.

The Consumer Health Resource Information Service Project is a faith-based pilot initiative designed by the Medical Education and Outreach group of the Oak Ridge Associated Universities, Oak Ridge, Tennessee. The purpose of the project is to address minority health disparities through community level intervention and prevention measures in six predominantly African American churches in the inner city of Knoxville. The project has provided health education, screenings, and services, and each church is equipped with a computer workstation and parish nurses trained to access medical information on the Internet.

The mission of the Environmental Health Information Outreach Program is to enhance the capacity of minority serving academic institutions to reduce health disparities through the access, use and delivery of environmental health information on their campuses and in their communities. The Program's 2005 meeting included representation from 12 HBCUs, 3 tribal colleges and 3 Hispanic Serving Institutions.

The Native American Information Internship Program provides an opportunity for representatives from American Indian tribes, Native Alaskan villages, and the Native Hawaiian community to learn about the NLM and the NN/LM and to use that knowledge to improve access to health information in their communities. Support in 2005 included assistance to an intern from Papa Pla Lokahi to develop a culturally appropriate smoking cessation Web site and an interactive game-style tutorial. Also, one of our interns from Hawaii is applying GIS technology to a project looking at potential relations between environmental exposure and health disparities.

The Central American Network for Disaster and Health Information is a group of health science libraries and information centers working together to enhance local health and disaster information management capacities with a goal of contributing to disaster preparedness in the region. It is a partnership between the NLM, Pan American Health Organization, and United Nations International Strategy for Disaster Reduction. The Network consists of centers in Honduras, Nicaragua, El Salvador and Guatemala (with support from U.K. Department of International Development). Additional centers, located in Panama and Costa Rica, were added in 2005 with financial support from the European Community Humanitarian Office. The Network centers enable health

professionals, government agencies, and others in their countries to access vital information previously unavailable. These libraries have acquired the knowledge, skills, and resources that promote delivery of reliable information. There are now over 7,000 full-text documents available online.

The National Medical Association is a national professional and scientific organization representing the interests of more than 25,000 physicians and their patients of African descent. SIS continued its collaborated with the Association to conduct online database training at their six regional meetings held each year.

SIS exhibited at over 40 conferences in FY 2005. Several of these provided opportunities for presentations or workshops about NLM's information resources.

Research and Development Initiatives

To meet the mission of providing information on toxicology, environmental health, and targeted biomedical topics to the world, SIS has been developing new ways of presenting the world of hazardous chemicals in our environment to a wider audience.

ToxSeek is a meta-search engine that enables simultaneous searching of many different authoritative information resources on the Web. It provides integrated search results from targeted systems and displays related concepts. ToxSeek was publicly released at the end of FY 2005 and in FY 2006 will undergo evaluation and enhancement based on user feedback.

The World Library of Toxicology, Chemical Safety, and Environmental Health will provide a Web portal to global information resources in toxicology, chemical safety, and environmental health. The World Library is being developed by SIS staff, and will provide a cyberhome for an ongoing participatory project in which voluntary representatives from participating nations provide crucial input and feedback to assure credible and high-quality sources of information. The World Library has been populated with information resource sets from 20 countries; information from 25 other nations is being arranged. With support from NIH's Fogarty International Center, the World Library will be released in FY 2006.

A new database, **Drugs and Lactation (LACT)**, was under development in FY 2005. LACT provides information on drugs and other chemicals to which breastfeeding mothers may be exposed. It includes information on the levels of such substances in breast milk and infant blood, and the possible adverse effects in the nursing infant and includes links to other NLM databases. LACT is scheduled for public release in mid FY 2006.

ToxMystery, a prototype Web site for children between the ages of 7–10, was under development in FY 2005, and a beta version will be ready for release in FY 2006. It provides an animated game-like interface that includes finding potential chemical hazards in a home and includes fun sound effects. Focus groups and feedback from the targeted user community will drive enhancements to ToxMystery.

The goal of the Public Health Law Information Project (PHLIP) is to create in the public domain a searchable database of public health law information that will be not only a guide for nonspecialists, but also an excellent technical resource those who are specialists in the field. The prospective audience includes concerned citizens; nonspecialist attorneys faced with public health law issues; public health practitioners, academics and other researchers; government officials, legislators and other policy makers; and any other interested persons who do not currently have access to PHLIP or who seek a richer understanding of such information. In a pilot project, the state of Delaware, the Widener University School of Law, the Delaware Academy of Medicine and SIS are collaborating to produce a searchable database containing statutes, regulations, and other materials from Delaware that pertain to public health.

In an intragency collaboration, SIS and the DHHS Office of Public Health Emergency Preparedness agreed to develop a system for **Radiological Event Medical Management**. Intended for use by physicians and first responders, the system will include algorithm-based guidelines for evaluation and management of individuals exposed to radiation during accidental releases, use of radiological dispersion devices, and use of improvised nuclear devices.

In these and other initiatives, SIS continues to search for new ways to be responsive to user needs in acquiring and using toxicology and environmental health, HIV/AIDS, and other specialized information resources.

LISTER HILL NATIONAL CENTER FOR BIOMEDICAL COMMUNICATIONS

Donald W. King, M.D.
Acting Director

The Lister Hill National Center for Biomedical Communications (LHNCBC), established by a joint resolution of the Congress in 1968, is a research and development division of the NLM. Seeking to improve access to high quality biomedical information for individuals around the world, the Center continues its active research and development in support of NLM's mission. The Center conducts and supports R&D in the dissemination of high-quality imagery, medical language processing, high-speed access to biomedical information, intelligent database systems development, multimedia visualization, knowledge management, data mining and machine-assisted indexing. An external Board of Scientific Counselors (Appendix 3) meets biannually to review the Center's research projects and priorities. The most current information about Lister Hill Center research activities can be found at http://lhncbc.nlm.nih.gov/.

Lister Hill Center research staff are drawn from a variety of disciplines, including medicine, computer science, library and information science, linguistics, engineering, and education. Research projects are generally conducted by teams of individuals of varying backgrounds and often involve collaboration with other divisions of the NLM, other institutes at the NIH, and academic and industry partners. Staff regularly publish their research results in the medical informatics, computer and information science, and engineering communities. The Center is often visited by researchers from around the world. The Lister Hill Center is organized into five major components: Cognitive Science Branch, Communications Engineering Branch, Computer Science Branch, Audiovisual Program Development Branch, and the Office of High Performance Computing and Communications.

Image Processing

The Lister Hill Center performs extensive research and development in the capture, storage, processing, retrieval, transmission, and display of biomedical documents and medical imagery. Areas of active investigation include image compression, image enhancement, image recognition and understanding, image transmission, and user interface design.

WebMIRS

Developed several years ago and still in active use, the Web-based Medical Information Retrieval System (WebMIRS) continues to provide access to images and text from nationwide surveys conducted by the National Center for Health Statistics. As described in previous annual reports, this Java application allows remote users to access data from the National Heath and Nutrition Examination Surveys II and III (NHANES II and III), carried out during the years 1976–1980 and 1988–1994, respectively. At the current time there are about 400 users of WebMIRS in 44 countries.

An important new tool, Multimedia Database Tool (MDT), will serve as the next generation WebMIRS system. The MDT will provide a software framework for the incorporation of new text/image databases in a much more general way than the current WebMIRS, and new features for the database end user that extend current WebMIRS capabilities. The specific framework that has been designed has the goal of accommodating new sets of text and images under a very flexible database schema and GUI approach. New features being incorporated for end users of the system include support for multiple levels of system privileges for users and capability for users at authorized levels to make new data entries into database fields.

Content-Based Image Retrieval (CBIR)

The goals of this project are to research and implement the latest technological approaches for indexing and retrieving biomedical images by direct use of image data and associated text. Our emphasis is on two-dimensional images, primarily the NHANES spine images, using shape methods on vertebrae in the images, and on NCI cervigrams, using color and texture methods to differentially identify tissue regions and tissue characteristics within these images. We also focus on development of effective CBIR methods to be incorporated into our multimedia database programs (such as the MDT) or into separate, prototype systems for use by the biomedical research and/or clinical communities.

The range of CBIR activities include: the development of integrated CBIR capability for retrieval by shape and partial shape; segmentation and truth data collection for the x-ray images; development of relevance feedback methods for shape retrieval for the spine x-rays; refinement of the Live Wire Segmentation technique; collection of vertebral segmentation data by medical experts; addressing problems of brightness removal and illumination correction in the cervigram images; development of Level Sets Segmentation methods for the spine x-rays; and the development of an integrated shape segmentation system.

A prototype CBIR client/server application was also developed. This application allows remote users to submit shape queries to shapes that have been indexed with the shape space indexing method developed by our Yale collaborators. The system incorporates a Java image server that delivers images using Texas Tech HVSQ compression. It is a set of modules developed in MATLAB and Java and is CEB's first demonstrable prototype for CBIR over the Web.

MARS

Medical Article Records System (MARS) progressed sufficiently to enable NLM to discontinue manual keyboarding entirely. A key element in allowing NLM to eliminate its keyboarding contract is the capability designed in MARS to accommodate foreign language journals (that account for 11 percent of MEDLINE citations). This requirement introduced new rules to extract vernacular titles (required in Roman script languages but not in others), and process the second pages of articles (to accommodate abstracts that spill over to a second page). These goals have been achieved by our FLEX software suite that is incorporated in several MARS workstations.

For some years MARS has evolved through several generations of increasing capability. Its core engine consists of daemons based on heuristic rule-based algorithms that use geometric and contextual features derived from OCR output to automatically segment scanned pages of journal articles, assign logical labels to these zones, and to reformat zone contents to adhere to MEDLINE conventions. About a quarter of the total citations in MEDLINE now are created by MARS, the rest come in as XML-tagged data directly from publishers.

WebMARS was developed to complement MARS by enabling the extraction of bibliographic citations from online journals. Since a majority of citations now come directly from publishers in XML format, WebMARS functions have been used to develop two other systems that will serve to increase the efficiency of creating citations for MEDLINE so that the expected doubling of the citation rate in a few years can be accommodated through automation, a goal of NLM's Indexing 2015 Initiative. One of the advantages of WebMARS is that all of the bibliographic data contained in the online article may be extracted.

Its first component, Publisher Data Review (PDR), will provide operators data missing from the XML citations sent in directly by publishers (such as databank accession numbers, NIH grant numbers, funding sources, and PubMed IDs of commented articles) thereby reducing the burden on operators in creating citations for MEDLINE. In addition, incorrect data sent in by the publishers can be corrected by PDR.

Currently, this is a labor-intensive process since the operators perform these functions manually by looking through an entire article to find these items, and then keying them in.

An initial version of the PDR system was completed after testing the software to handle databank accession numbers and grant numbers. Currently efforts are under way to discover discrepancies between those detected and actual data in MEDLINE. In collaboration with the Indexing Section, rules are being developed to eliminate these discrepancies, so that PDR can reliably provide operators data missing from the XML citations sent in directly by publishers. PDR will also correct incorrect data sent in by publishers. These features will help reduce the manual effort in creating citations for MEDLINE.

The second component, the WebMARS Assisted Indexing system, is for the indexers; it will help them search for terms in an article that correspond to biomedical terms in a predefined list. Again, indexers currently have to read through the entire article to confirm the occurrence of these terms, a labor-intensive process. The system will automatically search through the text and highlight these terms for the indexer to simply confirm and select, thereby reducing manual effort.

AnatQuest: A Window into the Visible Human

The goal of this project is to bring the high resolution Visible Human images to the lay public both directly as well as by linking text documents received from Web sources to relevant anatomic objects. This is achieved by two systems: AnatQuest and TILE.

AnatQuest is a Web-mediated system designed to provide widespread access to the Visible Human images for a broad range of users, including the lay public who are frequently limited to low speed Internet connections. This system is based on a 3-tier architecture in which the first tier consists of Java applets for displaying thumbnail images of the Visible Human Male, from which detailed (full-resolution) views are accessed. The second tier is a set of servlets that process user requests and compress the requested images prior to shipment back to the user. The third tier is the object-oriented database of high resolution VHP images and rendered 3D anatomic objects. Low bandwidth connections are accommodated by a combination of adjustable viewing areas and image compression done on the fly as images are requested. Users may zoom and navigate through the images. Since its release in 2003, it has averaged about 60,000 hits per month, about 5 times the number of hits for the AnatLine system developed in an earlier project intended mainly for the scientific and visualization communities.

Recently, to improve access to VHP images through common search engines, e.g., Google, the image labels were displayed below each thumbnail. This strategy has enabled Web crawlers to index the labels, thereby allowing the public direct access to the images through search engines, in excess of 10,000 hits in the first month after release.

TILE (Text to Image Linking Engine) is designed to transparently link the print library of functional-physiological knowledge with the image library of structural-anatomic knowledge into a single, unified resource for health information, a long term NLM goal. We interpret this goal as adding value to text resources such as PubMed and MedlinePlus by linking to anatomic images. A modular prototype TILE system is being developed to serve as a testbed to investigate the functions needed to accomplish this linkage.

Engineering Laboratories and Resources

The R&D relies on laboratories designed, equipped and maintained by the Branch, as well as content resources that support research as follows:

- Image Processing Laboratory. The CEB Image Processing Lab is equipped with a variety of high end servers, workstations and storage devices connected by a mix of 100 and 1000 Mb/s Ethernet. The laboratory supports the investigation of image processing techniques for both grayscale and color biomedical imagery at high resolution.
- Document Imaging Laboratory. This laboratory supports DocView, MARS and other projects involving document imaging. Housed in this laboratory are advanced systems to electro-optically capture the digital images of documents, and subsystems to perform image enhancement, segmentation, compression, OCR and storage on high density magnetic and optical disk media.
- Document Image Analysis Test Facility. Developed by the Communications Engineering Branch, this off-campus facility houses high-end workstations and servers that constitute the MARS production system.

Ground Truth Data for Document Image Analysis

For research in document image analysis and understanding techniques by the computer science and informatics communities, we provide a database named Medical Article Records Groundtruth (MARG). The data consists of over 1,000 bitmapped images of the first pages of articles from biomedical journals indexed in MEDLINE falling into 9 layout types encountered in MARS production. Included in addition to the page images are the corresponding segmented and labeled zones, OCR-converted and operator-verified data at the

zone, line, word and character levels, all in XML format. Also available from this Web site (marg.nlm.nih.gov) is Rover, an analytic tool that may be used to compare the results of a researcher's program with the ground truth data. The MARG server has had over 6,300 unique IP visits from 92 countries.

The Visible Human Project

The Visible Human Project (VHP) image data sets are designed to serve as a common reference for the study of human anatomy, as a set of common public domain data for testing medical imaging algorithms, and as a test bed and model for the construction of image libraries that can be accessed through networks. The Visible Human data sets are available through a free license agreement with the NLM. They are distributed to licensees over the Internet at no cost; and on DAT tape for a duplication fee. The data sets are being applied to a wide range of educational, diagnostic, treatment planning, virtual reality, virtual surgeries, artistic, mathematical, legal and industrial uses by over 2100 licensees in 49 countries. The Visible Human Project has been featured in over 850 newspaper articles, news and science magazines, and radio and television programs worldwide.

The Insight Toolkit (ITK), an R&D initiative under the Visible Human Project, continues to grow and influence the medical community. It is now in its fifth year with a recent ITK 2.2 software release. ITK makes available a variety of open source image processing algorithms for computing segmentation and registration of high dimensional medical data on a variety of hardware platforms. Platforms currently supported are PCs running Visual C++, Sun Workstations running the GNU C++ compiler, SGI workstations, Linux based systems and Mac OS-X. A consortium of university and commercial groups is executing this work.

The Visible Human Project has extended awards to extend the software infrastructure into clinical and research applications through the introduction of database management tools, workbenches for tumor volume measurement for possible use in clinical trials, and the sponsorship of Web portals for sharing research data and publications. Non-funded researchers from across the world are now testing, developing and contributing to ITK in over 38 countries. ITK users are represented by over 1000 active subscribers to the global mailing list for the project.

NLM has attained its primary goal of creating a strong, usable, public, open-source software infrastructure to support medical imaging research. The National Alliance of Medical Image Computing, an NIH Roadmap National Center for Biomedical Computing, has adopted ITK and its software engineering practices as part of its engineering infrastructure. The Alliance is

currently using medical imaging techniques to study the physiological sources of schizophrenia and other mental disorders. During FY 2005, the Alliance has held nine user-training workshops across the country, as well as an advanced ITK workshop in Lausanne, Switzerland. Other NIH ICs and their sponsored projects are not only taking up ITK, but they are helping to spread its influence and applying the toolkit to real-world problems. In addition to the NIH sponsored projects, a new non-profit organization, the Insight Software Consortium, was formed in FY 2005 to undertake long-term advocacy for ITK and its concerns.

3D Informatics

The Office of High Performance Computing and Communications 3D Informatics Program has expanded in-house research efforts around problems encountered in the world of 3-dimensional and higher-dimensional, time-varying imaging. One of our most intense efforts is a project to create PLAWARe (Programmable Layered Architecture With Artistic Rendering) a software framework for artistic and non-photorealistic rendering of digital models. This entails the design of a layered, software architecture for implementing medical illustration techniques using computer graphics technologies. In FY 2005, the project produced a working proof-of-concept implementation and submitted the results of the research for publication.

We have extended and enhanced our pilot project for creating the framework for an archive of volume image data, the National Online Volumetric Archive (NOVA). In FY 2005, we added a diverse collection from the Mayo Clinic including multi-modal imaging data from across a range of anatomy, physiology, pathology, and even animal models. Also in FY 2005, we took some early steps toward management of a cancer imaging collection of digital colonography data that includes the source x-ray CT patient, the radiologist's reports, the results of a virtual colonoscopy scan, images and reports from an endoscopic colonoscopy exam, and pathology reports of any polyps removed during the procedure.

Visible Human Educational Collaboration

The NLM has endorsed the acquisition and application development through the anatomical and Next Generation Internet programs which has used raw visible human data in the past decade, to pass from a data to knowledge stage. This data acquisition led to the development of educational applications affecting the discipline of gross anatomy, by contracts awarded to develop knowledge-based outcomes that can be applied to basic science curricula.

As evidenced by requests for invited sessions to professional anatomy societies, most educational medical and allied health institutions are eager to see continued development of Visible Human applications. The timing of creating educational collaboratories is also important due to the compression of curriculum time in the anatomical sciences, and the merging of some anatomy courses with physiology. The American Association of Anatomists, the American Association of Clinical Anatomists, and the Human Anatomy and Physiology Society strongly endorse the Visible Human's use in basic science curricula.

The purpose of the "Collaboratory" is to develop academic sharing mechanisms (across advanced broadband networks) for application software directed to basic science curriculum insertion in dental, medical, or allied health school settings, in the discipline of gross and microscopic anatomy. The goal of the program is to make available software applications based on Visible Human Project data to Collaboratory member institutions through multicasting or a similar networked broadcasting arrangement. This requires the planned integration of the applications into existing curriculum schedules. Therefore, a demonstration of this type of networked software applications, developed by one institution, can be linked by networking to other institutions of the Collaboratory. The Collaboratory approach involves image processing, information systems, infrastructure research, multimedia visualization, and training. The process becomes a multidisciplinary effort to take mature digital science program areas and focus the outcomes to creative public use in professional schools worldwide.

Information Systems

The Lister Hill Center performs extensive research in developing advanced computer technologies to facilitate the access, storage, and retrieval of biomedical information.

Consumer Health Informatics Research

Consumer Health Informatics research projects explore the needs, information seeking behavior, and cognitive strategies of health care consumers. Their principal goal is to apply medical informatics and information technologies to study ways to develop, organize, integrate, and deliver accessible health information to the members of the public at all levels of health literacy. These projects include the ClinicalTrials.gov and Genetics Home Reference Web sites and the Consumer Health Information Seeking research initiative.

ClinicalTrials.gov provides members of the public with comprehensive information about all types of clinical research studies, both interventional and

observational. The site has over 23,000 protocol records sponsored by the U.S. federal government, pharmaceutical industry, academic and international organizations in all 50 states and in over 120 countries. Some 47 percent of the trials listed are open to recruitment; the remainder are closed to recruitment or completed. ClinicalTrials.gov receives over 7 million page views per month and hosts approximately 20,000 visitors daily. Data are submitted by over 1,440 study sponsors through a Web-based Protocol Registration System (PRS) which allows providers to maintain and validate information about their trials.

ClinicalTrials.gov was involved in developing and implementing new standards to promote transparency in clinical research through trial registration. Concurrent with the release of the International Committee of Medical Journal Editors (ICMJE) statement requiring trial registration as a condition for publication in fall 2004, ClinicalTrials.gov expanded its scope to include trials sponsored by members of the international research community. ClinicalTrials.gov also added several new key fields specified by the World Health Organization minimum data set, thereby allowing sponsors to describe their research plans more fully. As a result of these new initiatives, over 9,000 new registrations were received over a five-month period beginning in May 2005.

The Genetics Home Reference (GHR) provides basic information about genetic conditions and the genes or chromosomes responsible for those conditions. Created for affected individuals, their family members, and the public, the site currently includes more than 170 condition summaries and more than 250 gene summaries, which are added at a rate of about 14 new summaries per month. GHR's content expanded to include all 29 conditions recommended by the Health Resources and Services Administration (HRSA) for inclusion in state newborn screening programs. In addition, GHR added a "spotlight" feature that draws attention to important initiatives related to genetic health that may be of interest to the public.

In the two years since its launch, the site's usage has increased nearly tenfold. GHR continues to receive favorable comments from healthcare providers, patients, family members, educators, and librarians. Recently, online media outlets such as Forbes.com, CNN, and Nature.com have linked to GHR to provide background material for genetics-related stories. A usability survey among the membership of the Genetic Alliance indicated that the site is easily navigable and the content is credible and understandable.

The Consumer Health Information Seeking initiative focuses on understanding and improving access to online health information. One project explores the search and navigation behavior of consumers using health information systems. Another project investigates methods for developing readability assessment metrics to evaluate health-related text intended for consumers of varying health literacy. A third project examines different approaches for using queries in one language (e.g., Spanish) to retrieve relevant documents in another language (e.g., English) to support access to health information for the Spanish-speaking community. A prototype system for providing basic information about clinical trials in Spanish is undergoing usability testing. Finally, the consumer health vocabularies project focuses on mapping words and phrases commonly used by consumers to technical medical terms and concepts.

Digital Library Research

Digital Library Research encompasses all aspects of creating, disseminating and preserving digital collections, developing metadata standards, applying emerging technologies and formats, and resolving copyright and legal issues. Research issues currently in focus are long-term preservation of digital archives, innovative methods for creating and accessing digital library collections, and developing modular and open information environments. Investigations concerning interoperability among digital library systems, the role of well-structured metadata, and varying "points of view" on the same underlying data set are also being pursued. The prime example of these activities is the Profiles in Science project. It uses innovative digital technology to make available the manuscript collections of prominent biomedical researchers, medical practitioners, and those fostering science and health.

The content Profiles in Science was created in collaboration with NLM's History of Medicine Division, which processes and stores the physical collections, including books, journal volumes, pamphlets, diaries, letters, manuscripts, photographs, audio tapes and other audiovisual resources. The collections of Francis Crick, Albert Szent-Györgyi, and Salvador E. Luria were added this year, bringing the number of scientists to sixteen. An additional 4,500 pages in 600 documents were also added to existing Profiles in Science collections. The 1964–2000 Reports of the Surgeon General, the history of the Regional Medical Programs, and Visual Culture and Health Posters are also available on Profiles in Science. Profiles in Science was recognized by ScientificAmerican.com as one of the top 50 science and technology Web sites and was highlighted in Science Magazine's "Best of the Web" each time a new collection was added.

During this fiscal year, several research projects continued to enhance the effectiveness of Profiles in Science. New scanning hardware and software were installed, configured, documented, and

extensively tested. Protocols for digitizing collections at other institutions were fine tuned and put to the test by NLM and Tennessee State Library and Archives staff. Protocols for digitizing video were also developed and tested on the Szent-Györgyi video clips. Suitable test sets of Profiles in Science master TIFF images were identified and provided to others for automatic metadata extraction experimentation. Development of a new XML-based Web front end and transition to a new XML-based search engine, as well as automated testing and verification tools, continue to be pursued. Several extensive tests were performed to ensure that the transition from the current Web front end to the new XML-based version will be seamless.

Document Imaging for the Biomedical End-User

The goal of this research area is to apply document image processing and digital imaging techniques to document delivery and management, thereby addressing NLM's mission of providing document delivery to end users and libraries. An additional focus is to contribute to the bulk migration of documents for purposes of digital preservation, also part of the NLM mission. The active projects in this area are DocView, DocMorph, MyMorph and MyDelivery.

- DocView. This Windows-based client software, originally released in January 1998 and subsequently improved over several generations, has 17,099 users in 193 countries, an increase of more than 1,000 new users and 3 countries over last year. Reflecting the declining worldwide use of TIFF for distributing document images (compared with PDF), and the age of the software itself, the use of DocView is expected to decrease.
- MyDelivery. The goal of the client/server MyDelivery communications system, seen as a successor to DocView, is to enable reliable and secure delivery of very large (gigabyte-sized, such as digitized x-rays) files, and large numbers (hundreds) of attachments in a single delivery, especially over unreliable links such as wireless networks. Recent developments have been in six critical areas: Log Shipping, Network Load Balancing, operation of the client behind proxy servers, operation of the client on all target operating systems, user interface, and managing the client.
- DocMorph and MyMorph. The DocMorph system continued to serve both browser-based users (over 12,500 to date: 2000 more than last year) and MyMorph users (over 5,200 users) this year. Of the more than 12,000 registered users, many are biomedical document delivery librarians. DocMorph allows the conversion of more than 50 different file formats to PDF, for instance, to

enable multi-platform delivery of documents. Also, by combining OCR with speech synthesis, DocMorph enables the visually impaired to use library information. Most users continue to use it to convert files to PDF to enable multi-platform delivery of documents. DocMorph is available at http://docmorph.nlm.nih.gov/docmorph.

By using Simple Object Access Protocol (SOAP) that combines XML with HTTP, MyMorph has been developed as a Web service that significantly improves the DocMorph function. MyMorph consists of Windows-based client software and modifications to DocMorph to accommodate SOAP. MyMorph significantly improves user productivity compared to the (conventional) use of DocMorph through a Web browser, particularly for users who need to convert large numbers of files to PDF. This is accomplished by reducing from hours to seconds the time required for users to interact with the software. The process of using MyMorph for converting image files to PDF has been integrated into many library document delivery operations worldwide.

MEDLINE Database on Tap (MDoT)

This project, previously known as PubMed on Tap, seeks to discover and implement systems and techniques to assist clinicians in quickly finding relevant, high quality information addressing clinical questions that arise at the point of care. An objective is to understand how to display data so that users can quickly find the most pertinent information while limited by the small screen and restricted bandwidth of handheld computers. Techniques were developed for display and navigation and for organizing the information. A prototype MDoT system was demonstrated.

As our primary method of discovery, we have developed a system that supports MEDLINE search and retrieval from a wireless, Internet-connected PDA. PDA client software for both Palm OS and Pocket PC OS have been developed and are freely available. Also, the MDoT Web site continues to be accessed at the rate of about 5,000 hits every month. This Web site provides information about the project as well as the software, and allows us to solicit feedback from our users and monitor aggregate user behavior.

Two studies were conducted to improve searching and user interaction. The first was a study of LHNCBC's experimental probabilistic search engine. It ranks results by relevance and was originally developed to support ClinicalTrials.gov. This investigation suggested that its search and ranking algorithms might be advantageous to MEDLINE searching at the point-of-care. A second study was conducted to assess the

value of the new e-spell function available from Entrez. Based on these studies, we designed a new Palm OS version 1.7 of the client with user interface and communications to support both an optional Auto Spell feature and an option to use the Essie search engine in lieu of the PubMed search engine.

Interactive Publications Research

This research effort was initiated in FY 2005 to create a comprehensive, self-contained and platform-independent multimedia document that is an "interactive publication." Following a study of existing open source formats and standards, a prototype document was created containing many media objects: text, dynamic tables and graphs, a microscopy video of cell evolution, an animated spine in Flash, digital x-rays, and clinical images (CT, MRI, ultrasound) following the DICOM standard. Both self-contained (embedded) and folder-type (linked) documents using all these media types were created in four formats: MS Word, Flash, HTML, and PDF. A comparison of these in terms of ease of use and development effort was done.

While using such a document, the reader is able to: (a) view any of these objects on the screen; (b) hyperlink from one object to another; (c) interact with the objects in the sense of exercising control over them (e.g., start and stop video); (d) and importantly, reuse the media content for analysis and presentation. On-going work will demonstrate and extend its usability and utility.

NLM Gateway

The NLM Gateway provides an easy to use "one stop" search method that allows users to issue simultaneous searches in a number of NLM information resources from a single interface. The current version interacts with five NLM search systems to provide results from 15 information resources. Changes to the underlying data structures or to the targeted search systems are carefully tracked and the Gateway modified accordingly.

While databases accessed by the NLM Gateway are regularly (sometimes even daily) updated, other resources incorporated into the Gateway itself are also regularly updated. New releases of the Unified Medical Language System Metathesaurus, the UMLS mapping file, the 2005 MeSH update, and Year End Processing were incorporated during the year as they became available.

A comprehensively redesigned NLM Gateway was released in April 2005. The new interface is cleaner, more straightforward and more understandable than the original, offering direct access to many functions and result sets with a single click of the mouse. The new system runs on Dell/Intel servers using the Linux operating system, making it possible to upgrade capacity at significantly lower cost than with the former Sun/Solaris servers.

As with the original, the new NLM Gateway is a meta-searching system, mapping user queries to appropriate search commands for multiple backend systems and issuing simultaneous searches in 15 databases. Results in terms of hit counts for all 15 databases are shown on one page. With one click, the user can display results from any of the 15 resources. While looking at those results, the user can with one click move to any of the other 14 result sets.

Digital Preservation Research

This project falls into two broad categories, one concerned with the preservation of documents and the other with video. In each area we focus on some of the key functions of an economical and robust digital preservation system. Two in particular are automated metadata extraction and file migration.

Extracting metadata automatically from the contents of material that need to be preserved, rather than relying on manual entry, is probably the only way large collections can be economically preserved. Techniques are also being developed that automate the migration of files in bulk. This is important for the conversion of files in formats that face obsolescence, largely because they are no longer supported by newer software and modern computers, and will be inaccessible as time passes.

For document preservation, a detailed design was done and a prototype System for Preservation of Electronic Resources or SPER was developed. SPER is a flexible, modular system that demonstrates key functions such as ingest, automated metadata extraction (AME) and bulk file migration. AME is implemented for the extraction of descriptive metadata from scanned and online journal articles as well as NLM's obsolete Web pages. The module for metadata extraction from Web pages is internal to SPER while those for TIFF and online journals are implemented via a SOAP interface, so that SPER can access a remote AME system and retrieve extracted metadata in XML format. In addition, a module for the extraction of technical metadata from TIFF file headers is implemented to include many of the items listed in the NISO Z39.87 standard for digital still images.

Research into video preservation focused on identifying an open file format such as Motion JPEG 2000 (MJ2) for archiving digitized video on disk media. Toward this end, a one-day invitational meeting was organized with about 50 archivists and technologists involved in the long term preservation of video and film. Participants considered the potential of lossless,

on-disk video storage in light of the "twilight of tape" as a cost-effective storage medium. Barriers to adopting lossless algorithms were identified, and specific directions to overcome them suggested. Also discussed were current video metadata standards, and recent work in automatic extraction of metadata from video.

Infrastructure Research

The Lister Hill Center performs and supports research in developing and advancing infrastructure capabilities such as high-speed networks, nomadic computing, network management, wireless access, and improving the quality of service, security, and data privacy.

Scalable Information Infrastructure (SII) Initiative

NLM's Scalable Information Infrastructure (SII) Initiative establishes testbed applications that demonstrate advanced network capabilities in health care, medical decision-making, public health, health education, or research within the broad agenda of the NLM. SII projects involve the use of testbed networks linking one or more of the following: hospitals, clinics, practitioners' offices, patients' homes, health professional schools, medical libraries, universities, research centers and laboratories, or public health authorities. Among the applications:

- Wireless Internet Information System for Medical Response in Disasters (WIISARD)—The University of California, San Diego in support of the SII initiative is developing an integrated software-hardware system designed to enhance the delivery of medical care by a local medical team at the site of mass casualty disasters.
- Advanced Network Infrastructure for Distributed Learning and Collaborative Research—The Haptic Audio Video Network for Education Technology (HAVnet) project is being conducted at the Stanford University School of Medicine. The project includes aspects of self-scaling technology, self-optimizing end-to-end network aware real-time middleware, wireless technology, and GIS.
- Project Sentinel Collaboratory—This project, a partnership involving units of Georgetown University and the Washington Hospital Center, is tasked with building and deploying a data-centric collaboratory to collect and analyze data from hospitals, clinics, weather services, satellite images of vegetation, mosquito collection, veterinary clinics and other sources in order to develop indicators and warnings of emerging threats to human health.
- National Multi-Protocol Ensemble for Self Scaling Systems for Health—The Boston Children's Hospital informatics program has developed a

secure XML medical record template for individuals, geographic information systems that can collect data on the use of these individual templates, and is negotiating with health care organizations for large scale applications (e.g., state immunization registries).

- Advanced Health and Disaster Aid Network—The Johns Hopkins Applied Physics Laboratory is developing a redundant network for the public health facilities of Montgomery County, Md., to be tested in mock public health emergency exercises with the anticipation of permanent installation if tests are successful.
- Advanced Network Infrastructure for Health and Disaster Management—The University of Alabama at Birmingham is utilizing emerging secure, high-speed wireless communication in the prehospital emergency medical services (EMS) and public safety information environment and integrating it with a hospital information system in an academic medical center and an independent EMS dispatching center.
- SMART—The emergency department at the Brigham and Women's Hospital in Boston will serve as the initial testbed for the SMART (Scalable Medical Alert and Response Technology), a system for patient tracking and monitoring from the emergency site that continues through transport, triage, and transfer from external sites to the health care facility within a health care facility.

Wireless PDA PubMed Searching

As part of the project, Information Technology for Low-resource Areas, a computer network was developed for health care organizations in low-resource environments. The network consisted of a central server, wireless access points, a dial-up connection to an Internet Service Provider, and featured the use of open-source software applications (Linux, Apache, PHP, MySQL). The wireless access points enabled mobile devices such as PDAs and wireless portable computers to disseminate administrative and health care information, and acquire reference resources and decision support tools through the Internet. An open-source software electronic medical record software, OpenEMR was initiated.

Telemedicine Initiatives

The HPCC Office participated in talks, demonstrations on state-of-the-art telemedicine, and e-Health projects in the Congress. The congressional Steering Committee on Telehealth and Healthcare Informatics was the sponsor as part of its work to address pressing healthcare issues including: advances in biosurveillance; efforts toward a

National Health Information Infrastructure; bringing information to the point of care; technologies toward effective disease management and for reducing medical errors; cost saving technologies; disaster and humanitarian assistance projects, and others. The NLM display featured wireless handheld access to PubMed.

A "Virtual Microscope" Website, http://erie.nlm.nih.gov/~slide2go was created to present the progress of the project. This site will archive images for medical education and telemedicine.

A Medical Image Database at the NLM (http://images.nlm.nih.gov) is also online. As individual images are viewed, a search form for MEDLINE/PubMed is automatically filled-out with the title of the image. A scanning digital microscope was acquired for the VM project.

The Telemedicine Information Exchange(TIE) is an NLM-sponsored Web-based resource of information about telemedicine and telemedicine related activities maintained by the Telemedicine Research Center, Portland, Oregon. During FY 2005 approximately 500 non-NLM bibliographic citations records were received at NLM in compliance with the contract.

Videoconferencing and Collaboration

Major upgrades to existing videoconferencing codecs were done in FY 2005 and new codecs were added. Demonstrations of streaming and wireless Webcasting were done and videoconferencing and Webcasting was employed routinely in the HPCC Office program. The adverse impacts of network security on the use of videoconferencing and collaboration technology were resolved by adding a new subnet outside the Lister Hill Center and NLM firewalls that is isolated from all other internal networks. Care has been taken to do regular security updates and virus checks on the machines connected to this subnet.

A distance learning program was conducted in collaboration with the Specialized Information Services, coordinator of NLM's Adopt-A-School Program. The program involved a series of presentations on medical topics and related health information resources and it was offered in conjunction with the Charles R. Drew University of Medicine and Science. The target audience for the program was students at the King Drew Medical Magnet High School in Los Angeles, one of the few medical magnet high schools for minorities in the U.S. Procedures were developed for recording the presentations and offering them later on demand by Webcast using the Collaboratory streaming server. The distance learning program was formally evaluated and found to be effective.

Work continues in experimenting with new codecs. Digital video compression technologies are being acquired for testing with members of the Internet2 community, since the digital video format is being considered as a compression format for the Access Grid. A beta version of Conference XP, an open source, experimental codec from Microsoft Research was installed on the Access Grid computer and tested with the Universities of New Mexico and Puerto Rico.

Language and Knowledge Processing

The Lister Hill Center conducts and supports research in language and knowledge processing to extract usable and meaningful information from biomedical text. This research covers terminology services, modeling and learning methods, medical ontologies, indexing initiative, and semantic knowledge representation.

Terminology Research and Services

LHNCBC research staff build and maintain the SPECIALIST Lexicon, a large syntactic lexicon of medical and general English that is released annually with the Unified Medical Language System Knowledge Sources. New lexical items are continually added using a lexicon-building tool; SPECIALIST contains over 295,000 records. Lexical access tools, including lexical variant generator, wordind, and norm, are distributed with the UMLS as are text processing tools which analyze documents into sections, sentences and phrases. The SPECIALIST lexicon, lexical tools and text processing tools are released as open source resources and available under an unrestrictive set of terms and conditions.

LexBuild, the lexicon-building tool, has been recently updated and moved to a Linux server. The SPECIALIST lexicon release tables are annually generated using the LexBuild tool. The SPECIALIST lexicon and tools are UTF-8 compliant and capable of dealing with non-ASCII characters. The lexical tools have been expanded so that they can be adapted to non-English language vocabularies. MetaMap Technology Transfer program (MMTx), the JAVA implementation of the MetaMap algorithm, is a major application of the SPECIALIST lexical and text tools. A part-of-speech tagger is being developed for use in MMTx that will exploit the SPECIALIST lexicon and will allow tagging of multi-word terms from the lexicon.

LHNCBC research staff also develop and maintain the UMLS Knowledge Source Server that provides Internet access to the UMLS knowledge sources through application programs and a user interface. The Knowledge Source Server is updated quarterly. Development has begun on a grid/Web services implementation of the Knowledge Source Server backend and an implementation of the user interface as a portal consisting of user chosen "portlets"

representing different parts and views of the UMLS data. These should result in a more flexible interface with both machines and human users.

The goal of the Terminology Server project is to provide tools and data to manage diverse medical vocabularies for diverse purposes. Over the past year, the project continued to provide customized data sets using the released versions of the UMLS to several projects such as ClinicalTrials.gov and Genetic Home Reference for use in their operational environments. An important function of the Terminology Server is to support the customization of terminologies from the UMLS and other sources. A number of internal tools were developed to handle the data customization needs of the projects identified above, which resulted in periodic releases of data sets containing customized data. One new set of tools handles the generation of English-Spanish translation tables for a Spanish version of ClinicalTrials.gov.

The project has continued to develop a set of tools that allow users to do their own data customization and management. In addition, the project will continue to integrate tools with existing applications and provide updates to the application data sets corresponding to the latest releases of UMLS data and other relevant data sets.

Modeling and Learning Methods

Digital information is at the center of the scientific endeavor and, if managed carefully, may provide a bridge for scientists and clinicians across disciplines. Today interdisciplinary research is severely hindered by language and terminology barriers such that scientific data and information generated in a field cannot be easily accessed, correctly interpreted and effectively used by other scientists.

The Modeling and Learning Methods project is aimed at developing computational learning methods to enable scientists to utilize cross-disciplinary information effectively. The multidisciplinary methods being developed in this project are rooted in biomedical informatics, machine learning, computational linguistics, and probabilistic knowledge modeling.

Recent research focused on effective information access by breaking the language barriers between the information and its potential users. Information retrieval, which currently is restricted to searching a corpus of text through a short list of keywords, is being enhanced with ontological, linguistic and probabilistic methods. Current prototypes can access, identify and retrieve biomedical information more effectively than the leading information retrieval engines, such as PubMed and Google.

A new probabilistic graphical modeling method called parameter interdependency networks has been developed. The method is substantially different from current probabilistic graphical models such as Bayesian networks, which are based on the concept of random variables. The new method on the other hand is based on a more fundamental concept called random events. The resulting method enables designers to model the system of interest (e.g., a biological process or a clinical problem) in the desired level of detail and provides a rich set of interactions between the model components.

Medical Ontology Research

While existing knowledge sources in the biomedical domain may be sufficient for information retrieval purposes, the organization of information in these resources is generally not suitable for reasoning. Automated inferencing requires the principled and consistent organization provided by ontologies. The objective of the Medical Ontology Research project is to develop methods whereby ontologies can be acquired from existing resources and validated against other knowledge sources. Although the UMLS is used as the primary source of medical knowledge, OpenGALEN, the Gene Ontology, and the Foundational Model of Anatomy are being explored as well.

During this fiscal year, the LHNCBC research team focused on the role of reference ontologies—including the Foundational Model of Anatomy and the Chemical Entities of Biological Interest—for ontology alignment and integration. Analogously, a mapping between two important clinical terminologies—SNOMED CT and ICD-9-CM—was derived automatically from knowledge extracted from the UMLS. Such automated mapping techniques are expected to play an important role in the exploitation of electronic patient records, and ultimately to contribute to health information technologies.

Work on ontology visualization was also pursued with RxNav, a publicly available, cross-platform application for visualizing drug information represented in RxNorm, one of the standards for use in U.S. federal government systems for the electronic exchange of clinical health information. Other research areas of this group include semantic similarity among genes computed from the various model organism databases annotated to the Gene Ontology. The application of Gene Ontology-driven similarity to the prediction of functional properties of genes and to the annotation of new genes was investigated. Methods were developed for converting ontologies from frame-based to description logic-based representations, with application to the Foundational Model of Anatomy.

The Indexing Initiative project continues to investigate techniques for the automatic selection of subject headings for use in both semi-automated and fully automated indexing environments at NLM. The project's goal is to facilitate the retrieval of biomedical information from databases such as MEDLINE. Team members have developed an indexing system, Medical Text Indexer (MTI), based on three fundamental indexing methodologies. The first of these calls on the MetaMap program to map citation text to concepts in the UMLS Metathesaurus. The second approach, the trigram phrase algorithm, uses character trigrams to match text to Metathesaurus concepts, while the third uses a variant of the PubMed related-articles algorithm to find MeSH headings by using the existing indexing terms of articles related to the input text. Results from the three methods are restricted to MeSH, if necessary, and combined into a ranked list of recommended indexing terms.

The MTI system is in regular use by NLM indexers, made available to them through the Data Creation and Maintenance System (DCMS). In addition, the indexing terms produced by MTI are being used as keywords to access collections of meeting abstracts via the NLM Gateway. Two recent efforts to improve MTI's accuracy are the extension of MTI input from titles/abstracts to the full text of articles and a word sense disambiguation facility to reduce MetaMap ambiguity.

Journal Descriptor Indexing (JDI)

The Journal Descriptor Indexing (JDI) project investigates a novel approach to fully automated indexing based on NLM's practice of maintaining a subject index to journal titles using a set of 122 MeSH terms known as JDs (journal descriptors) that correspond to biomedical specialties. For example, the *Journal of Pediatric Surgery* is indexed by the JDs Pediatrics and Surgery. JDI was used as a broad filter to extract from a ten-year MEDLINE text collection of 4.59 million records those likely to be of genomics interest (39 percent of the collection), as part of the NLM participation in the 2004 Text Retrieval Conference. Project staff also developed an algorithm used in a MeSH gene matcher program that contributed to the 2005 Text Retrieval Conference.

Unified Medical Language System

The mission, scope, and content of the Unified Medical Language System Metathesaurus continued to grow and evolve in FY 2005. During the fiscal year, regular production releases of the UMLS Knowledge Sources

were made. The number of concepts has increased over this time by 20 percent to more than 1.2 million. The number of names for concepts has increased by 17 percent to 5.5 million. There are more than 100 contributing source vocabularies.

The format and content of these biomedical vocabulary files vary widely. Without unifying standards or common tools, it is difficult to understand and use any single vocabulary, and far more difficult to integrate multiple combined vocabularies. The UMLS installation tool, MetamorphoSys, allows the selection of desired content from the Metathesaurus and writes the desired subset in Rich Release Format (RRF). The RRF contains additional information allowing exact attribution of the sources for all its information. This allows specific mappings between vocabularies, correct inclusion and exclusion of specific sources, and simultaneous representation of a consistent UMLS view along with each source's own view, which may differ.

MetamorphoSys now includes an integrated RRF browser that allows users to view their own Metathesaurus subsets in both Raw Data and Concept Report views. This addition means that any vocabulary in RRF may be reviewed, studied, or compared with views in other applications. This will make it much easier for users to make and then to see, understand, and verify their chosen Metathesaurus subsets in their own applications.

The Metathesaurus group has begun to promote the Rich Release Format both as a general vocabulary standard and as an input format for vocabulary submission to the UMLS. This process will make the nature and meaning of information from each vocabulary more explicit and allow the use of shared tools for validation, browsing and use, thus increasing interoperability. We are pleased to note that several other vocabulary providers have begun to distribute their vocabularies in RRF.

Semantic Knowledge Representation

Innovative applications for providing more effective access to biomedical information depend on reliable representation of the knowledge contained in text. The Semantic Knowledge Representation project develops programs that extract usable semantic information from biomedical text by building on existing NLM resources, including the UMLS knowledge sources and the natural language processing tools provided by the SPECIALIST system. Two programs in particular, MetaMap and SemRep, are being evaluated, enhanced, and applied to a variety of problems in biomedical informatics. MetaMap maps noun phrases in free text to concepts in the UMLS Metathesaurus, while SemRep uses the Semantic Network to determine relationships asserted between those concepts.

The MetaMap Technology Transfer program is an exportable, Java-based version of MetaMap that runs under Windows, Mac OS X or Unix/Linux and is provided as a resource to the bioinformatics community. MMTx allows users to exploit the UMLS MetamorphoSys program to exclude or reorder the Metathesaurus vocabularies that MMTx uses. Users can also create MMTx data files independent of the UMLS, and the inclusion of source code with each release allows additional control of processing.

The development of SemRep is based on viable strategies for effective natural language processing and underpins foundational investigations in biomedical information management. At the core of this research is enhancement of linguistic coverage. Recent additions include a mechanism for interpreting hypernymic propositions, and current work addresses arguments of nominalizations, comparative structures, and coordination of predicates. A modification of SemRep, called SemGen, is being developed for identifying and extracting semantic propositions on the causal interaction of genes and diseases from the research literature.

Multimedia Visualization

The Lister Hill Center performs extensive research and development in the capture, storage, processing, retrieval, transmission, and display of multimedia biomedical data. Multimedia products include high quality video, audio, imaging, and graphics materials.

Turning The Pages Information Systems

The Turning the Pages (TTP) project has two aims: (1) to design efficient methods to reformat the paper volumes in NLM's historic collection to photorealistic form, and (2) to extend these virtual books beyond their application as beautiful museum pieces to information systems that augment the material in the originals, including delivery over the Internet.

Originating as a collaboration with the British Library in producing two virtual books, Blackwell's 18th century *A Curious Herbal* and Vesalius' 16th century anatomy book, in TTP form, we have since made significant progress. The process consists of scanning the pages, enhancing these high quality color images by Adobe Photoshop, creating animated 3D wireframe models of the pages and book cover using Alias Maya (an innovation in our approach), run on a computer by Macromedia Director software, and displayed on a touchscreen monitor in an exhibit kiosk. The library patron may "touch and flip through" each of these books in an intuitive manner that evokes the feel of a "real" paper volume. Three additional books from NLM's historic collection have been added for a current

total of five books in TTP form: Paré's surgical treatise, Gesner's Animalium, possibly the earliest book on zoology, and Johannes de Ketham's Fasiculo de Medicina (1494).

In 2005 we collaborated with Library of Congress staff to create TTP at their institution. They were invited to a demonstration of our kiosk version of TTP, and a technical discussion of the steps required to create it. In addition to using their own resources and knowledge to accomplish the scanning, image enhancement and 3D modeling, they needed our templates for the final stage (to produce the software providing the interactivity). At our suggestion, they selected books in the life sciences: one by a Dutch surgeon who spent a decade with pirates in the Caribbean (1678), the other an encyclopedia of flora and fauna in the New World (1635). The first book was shown at an event at the Library of Congress in April 2005.

"The Library as Place" Flash Web Site

In FY 2005, the Audiovisual Production Development Branch developed a Flash Web site containing the entire NLM DVD program, "The Library as Place." The original DVD program was based on a two-day conference at NLM which examined the role of libraries in the digital age. Given the number of conferences and lectures held at the NLM, APDB initially chose to use DVD technology as a means for storing and delivering conference assets including video, audio, slides, Web links, textual documents associated with the conference and also make it highly searchable using visual and keyword search engines on the DVD. NLM distributed over 10,000 copies of the resulting Library as Place DVD. The Flash Web site will serve as a prototype for delivery of future NLM conferences and programs.

Virtual Dialogue Programs

APDB has been providing ongoing support to the NLM Visitors Center for a series of unique interactive multimedia programs, described as "Virtual Dialogues." These programs profile some of the scientific leaders of our time and are based on extensive recorded interviews with them. The virtual dialogue programs utilize speech recognition technology to allow a user, speaking into a microphone, to interview one of the many experts on a broad range of topics that each scientist has previously addressed in a carefully constructed video interview. The scientist, appearing as video on the computer screen, responds directly to the viewer in an informal, conversational style. Thus far, the series includes interviews with Marshall Nirenberg, Joshua Lederberg, Julius Axelrod, Edward Feigenbaum, and Donald Lindberg.

Changing the Face of Medicine—Local Legends Program and Web Site

In consultation with the Office of Communications and Public Liaison and in collaboration with the American Medical Women's Association, (AMWA), APDB has been working on the NLM/AMWA Local Legends program and Web site. The Local Legends Web site highlights congressionally nominated women physicians from 50 states as selected by a committee within the AMWA. Eleven video profiles have been completed for the Web site. APDB participated in all aspects of quality control to assure that the Local Legends Web site is fully compliant with Section 508 of the Rehabilitation Act by applying Web content accessibility guidelines to the entire site before deployment.

Additional DVDs were prepared in FY 2005 including LHNCBC Research Projects, Profiles in Science: Updates, UMLS, AMPA/NLM Annual Meeting, NLM Program Highlights 2005, InformationRx Press Event Highlights, and NLM Board of Regents presentations.

Training Opportunities
Working towards the future of biomedical informatics

research and development, the Lister Hill Center provides training and mentorship for individuals at various stages in their careers. The LHNCBC Informatics Training Program, ranging from a few months to more than a year, is available for visiting scientists and students. Each fellow is matched with a mentor from the research staff. At the end of the fellowship period, fellows prepare a final paper and make a formal presentation which is open to all interested members of the NLM and NIH community. In FY 2005, the Center provided training to 47 participants from 12 states and 8 countries.

The Center continues to offer an NIH Clinical Elective in Medical Informatics for third and fourth year medical and dental students. The elective provides an overview of the state-of-the-art medical informatics in a lecture series by Center research staff, and offers an opportunity for independent research under the mentorship of expert NIH researchers. The eight-week program, established in 2001, maintains its focus on diversity through participation in programs supporting minority students, including the Hispanic Association of Colleges and Universities and the National Association for Equal Opportunity in Higher Education summer internship programs.

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION

David J. Lipman, M.D.
Director

The National Center for Biotechnology Information (NCBI) was established in November 1988 by Public Law 100-607 as a division of the National Library of Medicine. The establishment of the NCBI by Congress reflected the important role information science and computer technology play in helping to elucidate and understand the molecular processes that control health and disease. Since the Center's inception in 1988, NCBI has established itself as a leading resource, both nationally and internationally, for molecular biology information.

NCBI is charged with providing access to public data and analysis tools for studying molecular biology information. Over the past 16 years, the ability to integrate vast amounts of complex and diverse biological information created the scientific discipline of bioinformatics. It is now almost impossible to think of an experimental strategy in biomedicine that does not involve some dependence on bioinformatics. At the core of this shift is the flood of genomic data, most notably gene sequence and mapping information. NCBI will meet the challenge of collection, organization, storage, analysis, and dissemination of scientific data by designing, developing, and distributing the tools, databases and technologies that will enable the gene discoveries of the 21st century.

The Center meets these goals by:
- Creating automated systems for storing and analyzing information about molecular biology and genetics;
- Performing research into advanced methods of computer-based information processing for analyzing the structure and function of biologically important molecules and compounds;
- Facilitating the use of databases and software by researchers and health care personnel; and
- Coordinating efforts to gather biotechnology information worldwide.

NCBI has a multidisciplinary staff of senior scientists, postdoctoral fellows, and support personnel. NCBI scientists have backgrounds in medicine, molecular biology, biochemistry, genetics, biophysics, structural biology, computer and information science, and mathematics. These multidisciplinary researchers conduct studies in computational biology as well as the application of this research to the development of public information resources.

NCBI programs are divided into three areas: (1) creation and distribution of databases to support the field of molecular biology; (2) basic research in computational molecular biology; and, (3) dissemination and support of molecular biology and bibliographic databases, software, and services. Within each of these areas, NCBI has established a network of national and international collaborations designed to facilitate scientific discovery.

GenBank—The NIH Sequence Database

GenBank® is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences. NCBI is responsible for all phases of GenBank production, support, and distribution, including timely and accurate processing of sequence records and biological review of both new sequence entries and updates to existing entries. Integrated retrieval tools allow seamless searching of the sequence data housed in GenBank and provide links to related sequences, bibliographic citations, and other related resources. Such features allow GenBank to serve as a critical research tool in the analysis and discovery of gene function as well as discoveries that lead to identification and cures for a number of diseases.

Important sources of data for GenBank are direct sequence submissions to NCBI from individual scientists and genome sequencing centers and substantial staff and resources are devoted to the analysis and curation of sequence data. NCBI produces GenBank from thousands of sequence records submitted directly from researchers and institutions prior to publication. Records submitted to NCBI's international collaborators, EMBL (European Molecular Biology Laboratory) at Hinxton Hall, UK and DDBJ (DNA Data Bank of Japan) at Mishima, are shared through an automated system of daily updates. Other cooperative arrangements, such as those with the U.S. Patent and Trademark Office for sequences from issued patents, augment the data collection effort and ensure the comprehensiveness of the database. The database is comprised of two divisions, traditional nucleotide sequences and Whole Genome Shotgun (WGS) sequences which are contigs (overlapping reads) from WGS projects that may contain annotations.

The traditional and WGS divisions of GenBank combined reached a total of 100 billion bases this year, a milestone for the database and its collaborators. In FY 2005, approximately 11 million sequences were added to the traditional GenBank division, and the base count rose from 40 billion in August 2004 to 50 billion in

August 2005. The WGS division added 6 million sequences and 25 billion base pairs in FY 2005.

GenBank indexers with specialized training in molecular biology create the GenBank records and apply rigorous quality control procedures to the data. NCBI taxonomists consult on taxonomic issues, and, as a final step, senior NCBI scientists review the records for accuracy of biological information. Improving the biological accuracy of submitted data as well as updating and correcting existing entries are high priorities for the GenBank team. New releases of GenBank are made available every two months; daily updates are made available via the Internet and the World Wide Web.

When scientists submit their sequence data to GenBank, they receive an "accession number." The accession number serves as a tracking device and allows the scientist to reference the sequence in a subsequent journal article. Sequence data submitted in advance of publication are maintained as confidential, if requested. NCBI is continuously developing new tools, and enhancing existing tools, to improve access to, and the utility of, the enormous amount of data stored in GenBank. Sequence data, both nucleotide and protein, are supplemented by pointers to corresponding PubMed bibliographic information, including abstracts and publishers' full-text documents as available. GenBank provides links to outside sources such as biological databases and sequencing centers. In addition to literature information, GenBank also provides links to related information in other Entrez databases. The availability of such links allows GenBank to serve as a key component in an integrated database system that offers researchers the capability to perform comprehensive and seamless searching across all related biological data on the NCBI Web site.

The Third Party Annotation (TPA), database created in conjunction with international counterparts EMBL and DDBJ, supports third party annotation of sequence data already available in public databases. Sequences in the TPA database are predicted or assembled from such sources as ESTs, genome data, and other unannotated sequences. Publication of the analysis in a peer-reviewed scientific journal is a requirement of this database. NCBI also accepts submissions from Whole Genome Shotgun sequencing projects. Annotations are allowed in these assemblies and records are updated as sequencing progresses and new assemblies are computed.

Improvement of NCBI's sequence submission software continues to be a high priority. Sequin, NCBI's stand-alone submission tool allows for updating as well as submission of GenBank sequences. The submission tool Sequin MacroSend allows submitters to upload a Sequin file from their computer directly to the GenBank indexing staff where their submission is immediately given a temporary identification number. Guides for specialized submissions such as genomes, batch sequences, and alignments are also available on the GenBank site. BankIt, another sequence submission software tool, is now in its eleventh year of use. Some of the improvements made to BankIt this year include an option for uploading an update file directly to NCBI. GenBank contains several types of sequence information, from relatively short Expressed Sequence Tags (ESTs) to assembled genomic sequences that are several hundred kilobases in length. EST data obtained through cDNA sequencing are critical to understanding gene function and therefore continue to be heavily represented in GenBank. The Genome Survey Sequences (GSS) division of GenBank contains sequences that are genomic in origin, rather than cDNA. The Sequence Tagged Site (STS) division consists of short sequences that are operationally unique in the genome and used to generate mapping reagents. Expanded and comprehensive STS information can be found in the UniSTS database.

Genome Resources

Integrated Genome Data

NCBI has developed a suite of genomic resources to support comprehensive analysis of the human genome, as well as the complete genomes of several model organisms. Specialized tools and databases have also been designed to facilitate researchers' use of data. NCBI maintains an expanding collection of specialized, yet integrated, database repositories that collectively capture and redistribute the biological relationships between genome sequences, expressed mRNAs and proteins, and individual sequence variations.

NCBI's Web resource, "Human Genome Resources," serves as a nexus for the collection and storage of diverse human data. This online guide provides centralized access to a full range of genome resources, including links to BLAST, dbSNP, RefSeq, Map Viewer, Gene, Homology Maps, UniGene, HomoloGene, and GEO. NCBI's Human Genome Sequencing site provides access to information on sequencing efforts and various other types of resources, such as chromosome-specific mapping information, cytogenetics, and TaxPlot for genome similarity plotting.

NCBI genome resource guides provide information on diverse organism-related resources from multiple centers including sequence, mapping, and clone information. The guides also provide easy navigation to organism-specific BLAST pages, and other NCBI resources. Genome resource guides for 23 organisms provide pointers and links to NCBI and outside resources such as documentation, maps and

sequence information, annotation projects, assembly updates, and comparative genomic sites.

Assembling and Annotating Genomes

NCBI is responsible for collecting, managing, and analyzing human genomic data generated from the sequencing and genome mapping initiatives of the public Human Genome Project. NCBI also plays a key role in assembling and annotating the human genome sequence. This resource is truly an international public sequencing effort due to the cooperation of scientists and sequencing centers from around the world. Human genome build 35 has been improved upon with increased annotation including access to alternate assemblies of the genome previously unavailable and reference sequences for two alternate haplotypes. NCBI's experience with curation of the human genome has benefited the annotation pipelines for other organisms as well.

A team of NCBI scientists is engaged in annotating, or characterizing, the biologically important areas of genomes. Genome builds are based on Gnomon, the gene prediction program developed by NCBI scientists. Gnomon uses RefSeq transcript sequences and EST alignments along with ab initio gene predictions, putting a greater emphasis on coding propensity and matches to existing proteins when predicting genes. As a result of the Gnomon program, the number of genes for human, mouse, and rat genome builds has decreased significantly, while the number of models identified as pseudogenes has increased. The number of human genes is now predicted as low as 20,000 versus earlier estimates of 35,000. NCBI's refined annotation pipeline allows annotation, in a single day, of microbial genomes submitted as whole genome shotgun sequences devoid of annotation. For example, the SARS coronavirus genome was annotated in one day.

Gene Identification

The Reference Sequence (RefSeq) database is a comprehensive, integrated, non-redundant set of sequences, including genomic DNA, gene transcript (RNA), and protein products for major research organisms. These standards serve as a basis for medical, functional, and diversity studies by providing a stable reference for gene identification and characterization, mutation analysis, expression studies, polymorphism discovery, and comparative analysis. In FY 2005, the NCBI RefSeq database grew by over 500,000 proteins and the most recent full release of all NCBI RefSeq records, Version 13, includes over 1.8 million proteins from 3,000 organisms.

NCBI is working with other groups to compare and evaluate genome annotation data and identify the set of proteins, as annotated on genomic sequence, which pass quality tests, and are consistently identified by different groups. The Consensus CoDing Sequence (CCDS) database is a collaborative project between NCBI, European Bioinformatics Institute (EBI), University of California, Santa Cruz, and Wellcome Trust Sanger Institute. The project is an effort to identify a core set of human protein coding regions that are consistently annotated and of high quality. Annotated genes included in CCDS are given a unique identifier similar to the GenBank system (accession.version). The initial CCDS dataset contains 14,795 coding sequences and 13,142 genes, representing more than half of the known human genes.

Entrez Gene is a significant step in offering a much larger scope of gene-specific data, better integration with other NCBI databases, and enhanced options for query and retrieval from the Entrez system. Entrez Gene integrates information about genes and gene features annotated on RefSeqs, collaborating model organism databases, making it easier for researchers to find and interpret gene-specific information. Gene replaced the LocusLink interface in March 2005. Currently, more than 2,900 taxa are represented in the Gene database with a total of about 1.3 million genes. Recent updates to the database include a more comprehensive navigation menu, more detailed displays, direct links to the new Splign tool, and improved links to RefSeq.

NCBI's Map Viewer continues to be the primary resource for visualization of large genomes. Increased standardization of map features better supports cross-species comparison and multiple-species queries. Maps from other sequencing centers are also available. Genes or markers of interest can be found by submitting a query against a whole genome, or by querying one chromosome at a time. The results table includes links to a chromosome graphical view where a gene or marker can be seen in the context of additional data. The Evidence Viewer is a feature that provides graphical biological evidence supporting a particular gene model and the Model Maker allows users to build a gene model using selected exons.

In FY 2005, NCBI continued to enhance its Map Viewer. The Map Viewer home page was reorganized for easier access to organisms of interest. Thirty-five organisms are currently represented in the Map Viewer. A multi-species view allowing maps from human, rat, and mouse to be displayed in tandem for detailed comparison was expanded to include chimpanzee.

As the number of sequenced genomes continues to grow, there is increasing interest in comparative analysis of genes from represented species.

The NCBI HomoloGene database of gene homologs performs such large-scale comparison automatically presenting reports that include statistics on inter-species sequence and protein domain conservation with links to the genome-wide views available in MapViewer, Entrez Gene, and gene expression information in UniGene. Over the past year, additional species were added to cover 18 model plant and animal genomes as well as the ability to view multiple species sequence alignments simultaneously.

UniGene is NCBI's system for automatically partitioning transcribed sequences into a non-redundant set of gene-oriented clusters. Each UniGene cluster contains sequences that represent a unique known or putative gene, as well as related information such as the tissue types in which the gene has been expressed, and map location. Curators are now able to classify cDNA libraries on attributes such as organ/tissue and developmental stage. Using this information, computational procedures for displaying gene expression profiles, identifying expression biases, and finding groups of genes with similar profiles has been developed.

Genetic Disease Information

Genes and Disease is a collection of articles designed to educate the lay public and students on how genes are inherited and cause disease and how an understanding of the human genome will contribute to improving diagnosis and treatment of disease. This collection, part of the NCBI Books site, contains descriptions for over 150 genetic diseases and links to databases and organizations for additional information. For each gene description there is a link to PubMed, the Online Mendelian Inheritance in Man database (OMIM), the Map Viewer, LocusLink, and BLink for related sequences.

OMIM is an electronic version of Dr. Victor McKusick's "Online Mendelian Inheritance in Man," a catalog of human genes and genetic disorders. The database, produced at Johns Hopkins School of Medicine, contains over 15,500 records. OMIM also contains two maps showing the cytogenetic location of disease genes. The "OMIM Morbid Map" is organized by disease, and the "OMIM Gene Map" is organized by chromosome. During the past year, the database was improved by adding two new categories to describe existing and new entries, thus making it easier, for example, to distinguish among records that describe genes only, diseases only, and diseases with a defined molecular basis.

Online Inheritance in Animals (OMIA) authored by Dr. Frank Nicholas, is a recent addition to the NCBI web site. OMIA is a database of genes, inherited disorders, and traits in animal species other than human and mouse. It contains textual information and references, as well as links to other relevant records from OMIM, PubMed, and Gene.

The GeneTests database produced at the University of Washington is now being supported, as is OMIM, by contract from NCBI. GeneTests is used more than 25,000 times a day by genetics counselors and physicians for its comprehensive genetic testing information and genetic disease descriptions. Data produced by this database has been released in the Books database under the title "GeneReviews" with almost 300 reviews.

Model Organism Genomes

In FY 2005, NCBI released Build 1.1 of the zebrafish reference sequence genome assembly based on Zv4 assembly from the Sanger Institute. The annotation is also available in NCBI's Map Viewer. Zebrafish is an excellent model to study human development and disease. The mouse genome was the first model organism available on the NCBI website. The mouse genome resource guide has links to mapping and BLAST pages as well as information on sequencing progress, sequencing centers, strain resources, and a monthly newsletter designed for the mouse research community. In FY 2005, NCBI mouse Build 34 was released and represents a fourth generation composite assembly. Rat genome Build 3 was released in FY 2005 with new maps including an assembly map, EST alignment maps for human, rat, and mouse.

Comparative Genome Data

Entrez Genomes contains records representing over 2,000 species including bacteria, archaea, and eukaryotes, complete microbial genomes, a number of viroids, mitochondria, a broad host range of plasmids, and over 1,500 viruses. Links to related resources include Plant Genomes Central, SARS CoronaVirus Resource, and the Influenza Virus Resource.

The Entrez Genome Project database is based on cellular organism-specific genomic information, including but not limited to genome sequencing such as whole genome shotgun or BAC ends sequencing projects, large scale EST and cDNA projects, and assembly and annotation projects. The database is designed around a hub-and-spoke model, with an organism comprising the hub, and individual projects the spokes. This design allows the collection of disparate data that all refer to a single organism, conveniently displayed for easy access with references to all subprojects.

Plant Genomes Central is an integrated, web-based portal to plant genomics data and tools. It

provides access to large-scale genomic and EST sequencing projects and high resolution mapping projects. The plant genomic effort has one technical hurdle relative to other genomic efforts: the range of plant genome size is very large extending from approximately the same size as the genome of many small animals to more than five times as large as the human genome.

The Viral Genomes website provides a convenient way to retrieve, view and analyze complete genomes of viruses and phages. This site now contains over 1,600 records for more than 1,200 different species.

Other Specialized Databases and Tools

The Influenza Virus Resource was created at NCBI with data obtained from GenBank and the National Institute of Allergy and Infectious Diseases (NIAID) Influenza Genome Sequencing Project. This project aims to produce "real time" sequence information during flu season to provide assistance in flu vaccination decisions. This resource will prove to be valuable due to the rapid evolution of flu viruses and will include sequence analysis tools for flu sequences as well as links to other resources on flu viruses. A Flu Dataset Explorer was produced to provide an interactive tool for preliminary analysis of protein sequences from the NCBI Influenza Sequence Database or from a user's own file.

The Gene Expression Omnibus, or GEO, is a high-throughput gene expression/molecular abundance data repository, as well as a curated, online resource for storage and retrieval of gene expression data. Currently, GEO contains over 55,000 accessioned objects. GEO Profiles, previously Entrez GEO, contains over 13 million expression profiles accounting for hundreds of millions of individual expression points. GEO DataSet (GDS) contains dataset definitions to facilitate identification of experiments of interest. At this time there are 1,000 curated experiments in the database. Graphical and text query tools for gene profiles and datasets have been developed. Multiple clustering methods are available as well as links to other resources such as HomoloGene and Entrez Gene.

NCBI is also engaged in large-scale analysis of genetic variation as part of the international Haplotype Map (HapMap) project, a consortium of scientists and funding agencies developing a public resource to find genes that are associated with human health and disease. The HapMap catalogs common genetic variants in humans, where they occur in the genome, and how they are distributed among people within populations and among populations in different parts of the world. The project currently is mapping over 5 million markers of genetic variation. With this fine-grained map,

researchers will be able to zero in on specific genomic regions to find the genes involved in particular disorders. With increased interest in expanding investigations into patterns of DNA sequence variation, NCBI is expected to play a prominent role in the transformation of genomic knowledge into the development of tools for disease diagnosis and therapeutic intervention.

The dbSNP database of genetic variation is a comprehensive catalog of common human polymorphisms for the international research community. dbSNP continues to experience rapid growth, containing over 21 million submissions of human data which have been processed and reduced to a non-redundant set of 10 million refSNP clusters. 28 other organisms are represented in the SNP database. dbSNP began accepting submissions for published haplotype data this year.

A highly specialized database of variation, the Database for the Major Histocompatability Complex (dbMHC) contains variations found only in alleles of the Major Histocompatability Complex, a highly variable array of genes that plays a critical role in determining the success of organ transplants and is largely responsible for an individual's susceptibility to infectious diseases. The dbMHC contains not only the hundreds of sequences of MHC alleles, but also data on typing kits used by academic, clinical, and industrial laboratories. The database is expected to grow rapidly, and now supports two major projects. The first is a survey of Human Leukocyte Antigen (HLA) allele frequency distributions in various populations, critical to the strategy of establishing and searching bone marrow donor registries as well as in studies of HLA-associated disease susceptibility. The second is an effort to collect HLA genotype and clinical outcome information on hematopoietic cell transplants performed worldwide. Support for new projects related to Type 1 Diabetes and rheumatoid arthritis is expected in the coming year.

The Gene Expression Nervous System Atlas (GENSAT) database contains images of the mouse brain at several different developmental stages using both in situ hybridization and transgenic mouse techniques, with a searchable set of gene expression annotations. The project aims to map the location of gene expression in the mouse brain and will shed light on the genetics of disorders that affect the central nervous system and on the brain's response to both natural and foreign chemicals. The project is sponsored by the National Institutes of Neurological Disorders and Stroke (NINDS). GENSAT is available for searching in the Entrez retrieval system with images indexed on a number of searchable fields such as gene symbol, gene name, and mouse age.

The Probe Database, launched this year, stores molecular probe data together with information on success or failure of the probes in different experimental contexts. Nucleic acid probes are molecules that complement a specific gene transcript or DNA sequence useful in gene silencing, genome mapping, and genome variation analysis. Probe is included as a database in the Entrez retrieval system with entry links available to experimental results and published articles.

RNA interference (RNAi) is an emerging technology for silencing specific genes that is proving to be of great utility as a research tool and may have important clinical applications. This year, NCBI launched a new RNAi resource to capture results of RNAi reagents and experimental results, such as extent of gene silencing and a variety of phenotypic observations. It is fully integrated with the Probe Database, where the actual RNAi reagents are stored.

PubChem

The PubChem project is a key component of the NIH Roadmap project in Molecular Libraries and Imaging. The PubChem database is designed to be a repository for small molecule data and the foundation for the massive amounts of bioactivity data that will be produced by NIH-sponsored chemical genomics centers. The PubChem database contains information on millions of small molecules including their structures, properties, and activities. Three databases containing specialized information, PubChem BioAssay, PubChem Compound, and PubChem Substance were added to the Entrez retrieval system in FY 2004. At the end of FY 2005, 4 million substance records and 3 million compound records were included in PubChem.

PubChem BioAssay allows users to examine descriptions of each assay's parameters and readouts, and links to 188,500 substances and 171,500 compounds. PubChem Compound searches unique chemical structures and validated chemical depiction information describing substances in PubChem Substance. PubChem Substance contains chemical substance records and associated information. Information was voluntarily contributed by over 35 academic, commercial, and government organizations including NLM's Specialized Information Services division.

Version 2 of PubChem, released in FY 2005, includes improvements such as redefining the compound database to allow users to better distinguish between compounds and substances. A ranking tool was developed to assist in indexing chemicals with multiple names. Increased linkage to MeSH, PubMed, and toxicology resources was also added and a streamlined automated deposition system is being developed. An extensive set of links connect PubChem data and link them to other Entrez databases. Many compounds have literature citations in PubMed as well as links to the proteins and/or genes representing a protein they bind to. Links between substances and compounds characterize chemical constituents. Links between substances and bioactivity indicate a substance was tested in a particular assay. Compound-compound links correspond to similarity relationships. Compounds are searchable by chemical structures, by chemical properties, and by bioactivity.

A new Open Mass Spectrometry Search Algorithm (OMSSA) was released this year as well. OMSSA helps users to identify MS/MS peptide spectra by searching libraries of known protein sequences. OMSSA scores significant hits with a probability score developed using classical hypothesis testing, the same statistical method used in BLAST.

Taxonomy

The NCBI Taxonomy project provides a standard classification system used by the international nucleotide and protein sequence databases. The Taxonomy database contains the names and lineages of greater than 165,000 organisms, both living and extinct, represented by at least one nucleotide or protein sequence in GenBank. The database adds approximately 19,000 new species per year. The Taxonomy group is continuing to support the development of the Barcode of Life project, which is attempting to obtain diagnostic 'barcode' sequences from every species of life on earth. The Taxonomy browser allows searches for information on an organism or taxi's lineage. Searches of the NCBI Taxonomy database may be made on the basis of whole, partial, or phonetically spelled organism names, with direct links to organisms commonly used in biological research. The Taxonomy system also provides a 'Common Tree' function that builds a tree for a selection of organisms or taxa.

Protein Structure

NCBI's Molecular Modeling DataBase (MMDB) is Entrez's 'Structure' database, a compilation of all the structures in the Protein Data Bank (PDB). PDB is a collection of all publicly available three-dimensional protein structures, nucleic acids, carbohydrates and a variety of other complexes experimentally determined by X-ray crystallography and NMR and is maintained by the Research Collaboratory for Structural Bioinformatics (RCSB) and the European Bioinformatics Institute. MMDB currently contains over 30,000 unique, experimentally determined 3D structure records. MMDB is updated monthly, with the source PBD data checked for consistency in the purported chemistry, sequence, and 3D coordinates.

In FY 2005, MMDB began functioning as a molecular modeling server, maintaining alignments relating protein sequences in Entrez to known three-dimensional structure. A new type of database cross-link to "related structures" presents graphical summaries that show regions of proteins, which have suitable structural templates for modeling. 3D visualization of the sequence-structure alignments, which imply the molecular models, can be used readily to infer their properties.

NCBI's three-dimensional structure viewer, Cn3D, provides easy interactive visualization of molecular protein structures from Entrez. Cn3D also serves as a visualization tool for sequences and sequence alignments. What distinguishes Cn3D is its ability to correlate structure and sequence information. Cn3D features custom labeling options, coloring by alignment conservation, and a variety of file export formats that together make Cn3D a powerful tool for structural analysis. Version 4.2 of the Cn3D structure visualization tool was recently developed to function as a multiple sequence- and structure-alignment editor.

The Conserved Domain Database (CDD) is the Entrez database of sequence alignments and profiles defining protein domains as recurrent evolutionary modules. Identification of conserved domains within a protein sequence is also available via the CD-search service, which is run by default for each protein BLAST search. The Conserved Domain Database annotation staff produces curated hierarchies of models related by descent from a common ancestor, representing the ancient evolutionary history of protein and domain families. The staff uses 3D structure information, phylogenetic analysis, NCBI Entrez's resources and the published literature, to enhance alignment quality, annotate functional sites, identify relevant links to PubMed and the NCBI Bookshelf, and update domain family summary descriptions to reflect available knowledge of molecular function.

VAST, or the Vector Alignment Search Tool, is a service that identifies similar protein three-dimensional structures of newly determined proteins. VAST compares new proteins to those in the MMDB/PDB database and computes a list of structure neighbors, or related structures, which allows a user to browse interactively, viewing superpositions and alignments in Cn3D.

Literature Databases

PubMed is a Web-based literature retrieval system developed by NCBI to provide access to citations and abstracts for biomedical science journal literature. PubMed is comprised of journals indexed in NLM's MEDLINE database as well as others beyond the scope of MEDLINE. It is the bibliographic component of NCBI's Entrez retrieval system and provides links to full-text journal articles at Web sites of participating publishers, as well as to other related Web resources. Full-text journal PubMed links have increased from 4,400 in September 2004 to 4,774 in September 2005. Usage of PubMed by the scientific and lay communities has also grown considerably since its introduction in 1997, with over 2.8 million searches per day.

The new My NCBI tool allows customization of NCBI Web services featuring an option to automatically update and e-mail search results from user saved searches replaced the Cubby. My NCBI includes additional features for storing an e-mail address, filtering search results, displaying LinkOut settings, document delivery service, and Outside Tool preferences. A spell check feature was added to PubMed that provides the user with choices of alternate spellings for search terms.

LinkOut is a feature of Entrez designed to provide users with links from PubMed and other Entrez databases to a wide variety of relevant Web-accessible online resources, including full-text publications, biological databases, consumer health information, research tools, and more. More than 1,780 organizations have supplied links to their Web sites, representing a 25% increase from last year. Sources include over 1,350 libraries, 220 full-text providers, and 219 providers of non-bibliographic resources including biological databases. Together they offer links to 34 million Entrez records. LinkOut resources received more than 21 million hits per month, a 40% increase from last year. Enhancements to the LinkOut program include a new and user-friendly LinkOut display, ability to display LinkOut holdings as a filter tab on Entrez search results to facilitate access to LinkOut resources, and automatic application of a library icon to citations with free full text links.

The NCBI Bookshelf gives users access to the full text of 44 textbooks in the clinical and research areas of biomedicine. Books may be searched directly or found through links in PubMed abstracts. In addition to textbooks from commercial publishers, the Bookshelf also includes monographs authored by NCBI, NLM, and NIH staff. Use of the Books database has increased six-fold in the past year and about two million book pages per month are downloaded by users.

New books added in FY 2005 include *Inflammatory Atherosclerosis: Characteristics of the Injurious Agent, Molecular Biology of the Cell (4th Ed.), Alternative Medicine and Rehabilitation, Parkinson's Disease: Diagnosis and Clinical Management* and the NCBI resources, NCBI Short Courses, and The Genetic Landscape of Diabetes. Other existing books and collections were updated and expanded, including the *Eurekah Bioscience Collection* with the addition of 211 chapters, the *HSTAT Collection*

with the addition of 90 evidence reports and technology assessments, and the *Historical Works: Medicine in the Americas Collection*. A new electronic format, PDA, is available for *HSTAT AHRQ Evidence Report Summaries and Medical Microbiology*.

PubMed Central (PMC) is a web-based repository of biomedical and life sciences journal literature providing free and unrestricted access to full-text journal articles. This repository is based on a natural integration with the existing PubMed biomedical literature database of indexed citations and abstracts. Currently, PMC contains over 200 life science journals. The number of full-text articles available through the PubMed Central (PMC) journal archive has grown more than 20 percent in the past year, to 375,000. The additions have come from newly published material as well as from digitizing back issues that previously were only available in printed form. Among the journals joining PMC in FY 2005 was the first journal from the UK Wellcome Trust's collaborative effort to add back issues of journals such as *Biochemical Journal* to the PMC archive.

Initial development of a packaged version of the PMC software for use by other archiving agencies was released in FY 2005. This portable PMC (pPMC) system enables collaborating international archiving centers to replicate the PMC archive and service, thus increasing the viability of the archive. pPMC will be tested in the UK, Italy, and South Africa in the coming months.

NIH's Public Access policy to publications resulting from NIH-funded research went into effect on May 2, 2005. NCBI designed and implemented the NIH Manuscript Submission system (NIHMS) and is responsible for archiving the manuscripts in the PubMed Central database. The system provides a quick and easy-to-use system for manuscript submission.

The BLAST Suite of Sequence Comparison Programs

Comparison, whether of morphological features or protein sequences, lies at the heart of biology. The introduction of BLAST in 1990 made it easier to rapidly scan huge sequence databases for similar sequences and to statistically evaluate the resulting matches. In a matter of seconds, BLAST compares a user's sequence with up to a million known sequences and determines the closest matches.

The BLAST suite of programs is continuously enhanced for effectiveness and ease of use. BLAST genome pages allow for convenient searching of an organism of choice. The BLAST sequence comparison server is one of NCBI's most heavily used services and its usage has grown at a pace reflecting the growth of GenBank. Each day more than 180,000 BLAST searches are performed, with users submitting their requests through server/client programs and the World Wide Web. Additional hardware and improvements in the BLAST code have enabled response times to decrease despite increases in the size of the database and number of users.

Several programming changes to BLAST queuing and calculation of final alignments have improved the turnaround time for answering a user's query and reduced the peak load on the formatting machines, allowing more searches with fewer resources. A new BLAST report formatter was made available, improving the presentation and value of results. Other enhancements include a new alignment style for closely related sequences as well as applying existing alignment styles to searches of translated nucleotides that were not previously supported. A new graphical viewer added the option of retrieving results in HTML format. This option makes it easier for users to store or even produce the results on their own computer and simplifies NCBI processing of formatting requests.

An electronic version of the polymerase chain reaction technique (e-PCR) was introduced to help identify sequence landmarks called Sequence Tagged Sites (STS) within a nucleotide sequence. E-PCR works by looking for matches to STS primer pairs with the orientation and spacing required to produce a PCR-amplified DNA sequence of the expected size. Both forward and reverse PCR can be performed. Splign, the program used by NCBI to align gene transcripts to genomic sequences during the gene prediction process, was introduced on the NCBI website in FY 2005. Splign uses BLAST hits to identify possible locations of genes and their duplications on genomic sequences and then to speed up the core dynamic programming.

Other comparison tools include TaxPlot, a tool used for the comparison of a reference genome to two others on the basis of the sequences of the proteins they encode. TaxPlot uses pre-computed BLAST results and covers almost 50 eukaryotic and more than 320 bacterial genomes. A related tool, GenePlot, generates plots of protein similarity for a pair of complete microbial genomes allowing a clear graphical visualization of deleted, transposed or inverted segments between the two. BLASTLink, or BLink, presents pre-computed BLAST alignments to database proteins, generating views of protein similarity that highlight taxonomic context, protein structure, or degree of functional annotation.

Database Access

Entrez Retrieval System

The major database retrieval system at NCBI, Entrez, was originally developed for searching nucleotide and

protein sequence databases and related MEDLINE citations. With Entrez, users can search gigabytes of sequence and literature data with techniques that are fast and easy to use. A key feature of the system is the concept of "neighboring," which permits a user to locate references or sequences which are related to a given citation or sequence. The ability to traverse the literature and molecular sequences via neighbors and links provides a very powerful and intuitive way of accessing the data. Entrez currently consists of 31 integrated databases providing information on sequences, taxonomy, genes, and literature. With Entrez Global Query, users can search all Entrez databases simultaneously in seconds and can find the number of hits displayed for each database on one page.

The Entrez Programming Utilities (E-Utilities) are tools that allow users to perform batch search and retrieval without using the NCBI Web interface. During FY 2005, support for new download formats was added for several of the Entrez databases such as Gene, SNP, and Taxonomy while a new E-Utility, "espell" was implemented to check spelling within user-generated Entrez queries and to offer suggestions for misspelled terms. Support for the Simple Object Oriented Protocol (SOAP) was also implemented, providing a structured XML-based method of access designed to simplify new application development by NCBI's users.

Other Network Services

Usage of NCBI's Web services continues to expand as more information and services are added. NCBI staff continued to make access and usage easier with improved documentation and tutorials. A Web usability group is addressing issues such as improving awareness of underutilized services, implementing a better and more consistent means to navigate the NCBI site, establishing a content management system, and evaluating user experience of all services.

The NCBI Web site provides an integrated approach to accessing all of NCBI's databases and services as well as general information about NCBI, its research, data submissions, and updates. At the end of FY 2005, NCBI's site was averaging over 75 million hits daily with over 1.5 million users. Because of the mission-critical nature of NCBI's computing platforms for PubMed, Entrez, BLAST, and other services, extensive system monitoring is performed. Based on measurements taken every 15 minutes from 50 ISP monitoring sites across the U.S. and overseas, the average time to load the NCBI home page is 0.82 seconds, an average PubMed search takes less than 2.5 seconds and availability has been better than 99.5 percent.

NCBI has a number of network services that provide programmatic access to several important NCBI

databases. A monitoring program was developed to make sure all of these services are responsive and producing correct information. The program quickly notifies relevant staff members if any service for which they are responsible becomes unavailable or starts producing unexpected or incorrect results. The detailed diagnostic information provided by the program has allowed coding bugs, configuration errors, and hidden dependencies to be detected and fixed rapidly, greatly increasing the reliability and utility of the services being monitored.

Computer Systems

Security continues to be an important priority for NCBI systems management staff. In September of 2004, NCBI network staff installed a redundant pair of Cisco PIX 535 firewalls enabling the relocation of firewall rules and filters from the core routers to these dedicated high-performance devices. Another security-driven change was the restriction of outside access to NCBI systems. All outside access is now forced through a collection of hardened "bastion" hosts. These systems require a secure ssh connection from the outside and only permit an ssh connection to internal hosts. Three such systems protect the Bethesda, Maryland site; two protect the collocation site in Virginia.

For the PubChem project, software toolkits from OpenEye Scientific Software and Xemistry GmbH were purchased. These toolkits allow PubChem to detect the similarity or identity of records of chemical structures obtained from multiple sources. The OpenEye libraries permit PubChem developers to take chemical structures with different rendering schemes and represent them in a common standardized format. In addition to the software toolkits, 2-CPU Linux web servers, 12 2-CPU Windows SQL Servers and two Network Appliance SAN clusters were added to support the PubChem project.

The NIHMS project, part of the Public Access initiative for NIH-supported scientific publications, also required a dedicated complement of hardware and software. Software was obtained for converting a variety of input formats submitted by authors to a standard PDF format. Three Windows servers, three SQL servers and four Linux web servers also support the NIHMS project.

In the past 12 months, the computing infrastructure that supports the NLM's flagship public service, PubMed, was substantially upgraded. Approximately 25 new 2-CPU PubMed web servers running Linux were put into service, as were 15 2-CPU Linux servers that support various internal PubMed processing requirements. In addition to the six Sybase database servers that support the PubMed Web servers, approximately 25 new Microsoft SQL Servers were

provisioned to support a wide variety of GenBank- and PubMed-related public service and internal production projects. A major initiative was undertaken to evaluate, benchmark and obtain storage-area network storage to replace directly attached storage for most, and eventually all, of the production database servers.

Another high-profile NCBI public service, the sequence similarity search service BLAST is also receiving a major hardware update. As part of NCBI's goal of refreshing computer technology on a 3-4 year cycle, the approximately 110 2-CPU BLAST servers, which are currently based on the 1.4 GHz Pentium 3 32-bit processors, are being replaced with 3.2 GHz or better 64 bit Pentium 4 processors. This project is presently about half-completed. 150 Windows desktop PCs were replaced with new models in the past year. A project was also initiated to deploy a major operating system upgrade to existing systems, replacing the Windows 2000 Professional operating system with Windows XP Professional.

Research

Research is at the core of NCBI's mission. The Computational Biology and Information Engineering Branches are the main research branches of NCBI, with the latter branch concentrating on applied Research and Development. Each Branch comprises a multidisciplinary team of scientists that carries out research on a broad range of fundamental problems in molecular biology by developing and applying mathematical, statistical, and other computational methods. Research conducted by NCBI investigators has strengthened applications and database work and has led to the development of many new theoretical and practical models that have opened doors to new areas of research.

NCBI's basic research group is within the Computational Biology Branch and consists of 75 senior scientists, staff scientists, research fellows, and postdoctoral fellows. Projects focus on new computer methods to facilitate the analysis of genome sequences and molecular sequence databases due to the rapid growth in large-scale sequencing efforts. Other projects focus on such techniques as the analysis of particular human disease genes and the genomes of several pathogenic bacteria, viruses and other parasitic organisms, in many instances, in collaboration with experimental laboratories. Another focus of research is the development of computer methods for analyzing and predicting macromolecular structure and function.

New areas of research include: development of novel amino acid substitution matrices for improved sensitivity of sequence alignment programs, evolutionary genetics, analysis of gene regulatory pathways, the development of new modeling tools for tumor DNA data, single nucleotide polymorphism data analysis, analysis of malaria genomes for vaccine development, evolutionary analysis of protein domains and comparative genomics, the evolution and spread of malarial drug resistance, and development of mathematical models of genome evolution. New databases are also being designed for data on conserved protein domains and mRNA expression experimental results. Staff continued collaboration with other NIH Institutes for sequence analysis, gene identification, and the analysis of experiments on gene expression. Collaboration was also continued with several institutes worldwide on genetic linkage analysis problems. New research projects were initiated in support of the new PubChem molecular libraries project.

A Board of Scientific Counselors, comprised of extramural scientists, meets twice a year to review the research activities of NCBI. The high caliber of the work of this group is evidenced by the number of peer-reviewed publications, over 100 publications this year with more in press. The staff participated in numerous presentations at various scientific meetings, and at universities worldwide. Presentations were also made to visiting delegations, oversight groups, and steering committees. In addition, staff gave presentations at NCBI's weekly Computational Biology Branch Lecture Series which also hosts outside speakers.

The NCBI Postdoctoral Fellows program is designed to provide training for doctoral graduates in a variety of fields including molecular, computational, and structural biology as well as graduates in other fields who elect to obtain additional training in computational biology. The NCBI uses the NIH Intramural Research Training Award Program and the Fogarty Visiting Fellow mechanisms to recruit for this program.

Outreach and Education

NCBI continues to expand its outreach and education programs to increase awareness of its myriad of public databases and specialized tools and services. Over the past year, NCBI staff presented at numerous scientific exhibits, seminars and workshops; sponsored a number of training courses, both lecture and "hands-on" courses; and published and distributed various forms of printed information.

Education: NCBI Courses

In response to an ever-increasing demand for education and training in the use of the increasing diversity of NCBI's products and services, the course, "A Field Guide to GenBank and NCBI Resources," was expanded and is taught at NIH and throughout the United States upon request. The course consists of a

three-hour lecture, a two-hour hands-on practicum, and optional one-on-one sessions. An expanded two-day course entitled, "Enhanced Field Guide" premiered in FY 2005 that provides more detailed coverage of NCBI databases and tools with an advanced hands-on session included. The eleven-member teaching staff presented 79 courses to over 5,000 people in the past year.

Education: Mini-Courses and Lecture Presentations

NCBI offers ten bioinformatics mini-courses at NIH and outside institutions to provide a practical introduction to various resources. The two-hour courses are both problem-based and resource-based and include a review and hands-on session. This year, 82 mini-courses were offered to over 3,000 participants.

Education: Technical Workshop Series

Debuting in July 2004, the PowerTools series consists of three courses lasting three days each. "NCBI Power Scripting" includes lectures and workshops on effectively using the NCBI Entrez Programming Utilities (eUtils) with scripts to automate search and retrieval operations across the entire suite of Entrez databases. The "NCBI 4-Pack" course provides information on practical applications of bioinformatics resources. "Programming with NCBI BLAST" presents effective uses of BLAST within scripts, creating and maintaining local BLAST databases, and setting up a BLAST web server. The courses, offered at NCBI, have been well-attended, most to capacity of 20 people per course.

Education: Bioinformatics Training

To help NIH researchers make optimal use of computer science and technology to address problems in biology and medicine, the NCBI has an intramural Core Bioinformatics Facility (CoreBio)—a network of bioinformatics specialists serving individual institutes within the NIH. Individual CoreBio Members are trained over a nine-week period in the use of bioinformatics tools provided to the research community by NCBI. The CoreBio Members, in turn, advise researchers within their respective Institutes as to the best methods for conducting their bioinformatics analyses. Information exchange among the CoreBio Members and the NCBI faculty is facilitated by regular meetings and e-mail forums.

CoreBio has trained representatives from 15 research Institutes at NIH, having conducted eight nine-week training programs, two in the past year since the program began in 2001. Twenty-five update sessions and two special topic sessions for the institute

representatives have also been held. One-on-one consultations are available on an ongoing basis for NIH scientists with NCBI faculty in the NCBI Learning Center, the NIH Library, and the NCI-Frederick Cancer Research and Development Center.

Education: Extramural Educational Collaborations

The educational collaboration program was established to train a network of bioinformatics support specialists who provide local educational and user support services for a wide range of users and needs. The university medical library is becoming a centralized point for providing these services at the local level, and members of the collaboration are based in institutions that are leading this trend.

The fourth, "NCBI Advanced Workshop for Bioinformatics Information Specialists," was held in FY2005. The collaborators and course alumni offer a variety of year-round services at their universities, including workshops on NCBI resources, individual research consults and support, and web portals. Many of the workshops are based directly on materials presented in the Advanced Workshop, thereby extending the impact of original materials. Together, the collaborators and course alumni form the growing Bioinformatics Support Network (BSN), a group supported by NCBI which has been established for the purpose of communication and continuing education among members.

In FY 2005, 12 new members joined the BSN as a result of the workshop, and two members from earlier years joined the group of educational collaborators. BSN members, and in particular Educollab members, are leaders in their field and have significant influence on the growing movement of medical libraries to establish high quality bioinformatics support programs. The regional training program for the three-day introductory course continued. Six of NCBI's educational collaborators served as regional instructors to present the course at five locations across the country. A total of 89 individuals attended these regional courses and the on-site course offered at NLM.

The purpose of both the introductory and advanced workshop, as well as the Educollab program, is to train the trainers, who then provide assistance with NCBI resources to thousands of end-users across the country. The Web-based materials, including newly added hands-on exercises, serve as a reference for those who have taken the courses. Because the exercises are based on frequently asked questions and actual user needs, they can also serve as a knowledge base for library staff that is assisting users in their academic communities.

Outreach: User Guides for NCBI Resources

NCBI has continued to develop a comprehensive list of fact sheets that outline the services and databases offered by NCBI. These fact sheets and guides are available for printing via the "About NCBI" site. In addition, a number of other informational and educational resources are available on the NCBI Web site. Links are available that discuss the fundamental principles of biomolecular research and underlying sequence similarity search tools. Interactive tutorials may be found for a number of databases and search and retrieval tools such as Entrez, PubMed, Structure, and BLAST. A comprehensive "NCBI Handbook" of 24 chapters has been incorporated into the Books database, as well as help documents for PubMed and Entrez, where they are continuously updated.

NCBI News is a quarterly newsletter designed to inform the scientific community about NCBI's current research activities, as well as the availability of new database and software services. The newsletter contains information on user services, announcements of new or updated services and available genomes, NCBI investigator profiles, and a bibliography of recent staff publications. In FY 2005, over 18,000 printed copies of the NCBI News were distributed quarterly. Access to the newsletter and its archives via the NCBI Web site has increased dramatically as more people have become aware of its availability online.

Biotechnology Information in the Future

Over the past few years, there has been an explosion in the volume of genomic data produced by the scientific community, most notably in the amount of whole genome, and gene sequence and mapping information. This is due in a large part to the interest generated by the human genome project and the follow-on projects involving the sequencing of whole genomes from model organisms and an array of other species. The commitment to providing the scientific community with both the resources and tools needed to fully explore this data as quickly as possible, as well as recent advances in molecular analysis technologies, promises that the exponential growth in genomic data will only increase. This reinforces the need to build and maintain a strong infrastructure of information support. NCBI, a leader in the fields of computational biology and bioinformatics, plays an active and collaborative role in deciphering the human, as well as other genomes and in developing state-of-the-art software and databases for the storage, analysis, and dissemination of data. The genomic information resources developed and disseminated thus far by NCBI investigators have contributed significantly to the advancement of the basic sciences and serve as a wellspring of new methods and approaches for applied research activities. The value of these resources will continue to grow, as NCBI is committed to the challenge of designing, developing, disseminating, and managing the tools and technologies enabling the gene discoveries that will significantly impact health in the 21st century.

EXTRAMURAL PROGRAMS

Milton Corn, M.D.
Associate Director

The NLM Extramural Programs Division (EP) continues to receive its budget under two different authorizing acts: the Medical Library Assistance Act (MLAA, unique to NLM), and Public Health Law 301 (covers all of NIH). The funds are expended mainly as grants-in-aid, and in some instances as contracts, to the extramural community in support of the Library's goals. Review and award procedures conform to NIH policies. The EP Web site at http://www.nlm.nih.gov/ep/funded.html lists grants awarded since 1997, with links to abstracts provided in the NIH CRISP (Computer Retrieval of Information on Scientific Projects) database.

EP issues grants in a broad variety of programs, all of which pertain to informatics and information management with the exception of the Scholarly Works program:

- Resource Grants for information management, often involving medical libraries;
- Training and fellowship grants for training informaticians and informationists;
- Research Grants in biomedical informatics, information science, and bioinformatics;
- Research Resource grants to support unique research resources tools for informatics and bioinformatics;
- Scholarly Works and Conference grants to enhance scientific and scholarly communication;
- SBIR (Small Business Innovation Research) / STTR (Small Business Technology Transfer Research) grants to support informatics innovations in small businesses; and
- Multi-Institute grant announcements and interagency collaborations.

In FY 2005, the NLM nominated a grantee, Dr. Kenneth Mandl, for a Presidential Early Career Award for Scientists and Engineers. NLM's nominee was one of those selected for the award.

Success Rates

In FY 2005, success rates for applications were significantly lower than in recent years. Only those applications with outstanding to excellent priority scores were given awards. The problem was particularly acute for three of EP's prime grant programs: informatics research grants (R01), Scholarly Works, and Knowledge Management/Applied Informatics grants, as shown in the table that follows.

Table 11

Average Priority Score, All Awarded Grants

Grant Type	*FY 2003*	*FY 2004*	*FY2005*
R01 Research	167	174	147
	(range 132-195)	(range 141-203)	(range 121-167)
Scholarly Works	153	167	142
	(range 126-177)	(range 126-183)	(range 118-176)
Knowledge Mgmt/	162	163	145
Applied Informatics	(range 137-188)	(range 129-201)	(range 113-166)

Success rates are expected to improve somewhat in FY 2006 because "out year" commitments to previously awarded grants, unusually high in FY 2005, will be less in the coming year. However, EP will not soon reach the 20–25 percent range seen during the years that the NIH budget was doubling. The combination of a relatively flat budget and a significant annual increase in applications received will probably result in awards for fewer than one in six.

The increasing number of awards EP makes through participation in NIH Roadmap, BISTI (Biomedical Information Science and Technology Initiative) and other multi-agency computing initiatives decreases the funds available for EP's own grant programs.

Because of the importance of training to NLM, individual informatics research training fellowships were given priority in fund allocation decisions for the MLAA budget. In FY 2006, it may be necessary to reduce support to individual fellowships in order to reallocate some funds to other important MLAA programs, such as career awards for new faculty and the knowledge management program. A reduction in individual fellowships does not impact the university informatics training programs, which carry out the great majority of the training supported by NLM.

Research Support for Biomedical Informatics and Bioinformatics (PHS 301)

Extramural research support is provided through a variety of grant mechanisms that support investigator-initiated research. EP's research grants support both basic and applied projects involving the applications of computers and telecommunication technology to health-related issues in clinical medicine and in research.

Research Grant Program
The R01 research grant program has two "branches," biomedical informatics and bioinformatics, both of

which are funded from PHS 301 funds. One hundred seven applications were reviewed for this program in FY 2005, compared to 48 applications in FY 2004. Ten awards were made. The average score of awarded applications was 147, compared to 174 in FY 2004. Twenty-five percent of R01 applications assigned to EP were reviewed in NIH's Center for Scientific Review, rather than NLM's Biomedical Library and Informatics Review Committee (BLIRC) . Three were deferred for funding in FY 2006.

Small Grant Program

To complement its R01 research grants, and to remain competitive with other NIH Institutes, in 2003 EP began offering the R03 small research grant, which provides relatively small amounts of support for "start-up" research projects. Thirty-six R03 grants were reviewed in FY 2005. Only four new grants were awarded. The average priority score of funded R03 grants was 163.

Informatics for Disaster Management

Since 2002, EP has had an R21 exploratory/developmental grant program exploring the application of informatics approaches in natural and man-made disasters. Fifteen new applications were reviewed in FY 2005, and one award was made. The success rate has continued to be poor for this program, as shown in the initial priority scores. The average score of Informatics for Disaster Management applications reviewed was 291; the average score of awarded applications was 150. In part the problem has been that many of the applicants propose systems rather than research. EP will consider refocusing the disaster management grant program from the R21 Research series into the Knowledge Management program as a better fit.

Exploratory/Developmental Grants

EP's R21 exploratory/developmental grant supports high risk/high yield projects, proof of concept, and work in new interdisciplinary areas. This grant mechanism is better suited to informatics/engineering proposals than are the standard R01 research grants which are judged in terms of hypothesis-based science. Thirty-eight applications were reviewed in FY 2005, compared to eight applications in FY 2004. Success rate was low with only four awards; the average score of awards was 168.

Resource Grants for Biomedical Informatics/ Bioinformatics

In August 2004, EP issued a P41 program announcement for support for scientific research resources. The program announcement was deactivated in FY 2005 for two reasons: (1) resource applications tend to be expensive and long-lived and constitute an important drain on funds available for research, and (2) our informatics interest is to focus more on the research questions than on sustaining service. Existing P41 awardees will be allowed to submit continuations, and fundable applications in the pipeline may be re-focused to R01 projects. Five new applications were reviewed in FY 2005, and one was funded. The average score of awards was 185.

Conference Grants

Support for conference and workshops is intended to provide relatively small amounts to scientific communities convening workshops and meetings in focused areas of biomedical informatics and bioinformatics. Of five applications reviewed in FY 2005 two were funded.

IAIMS Testing & Evaluation Grants

An updated program announcement was issued for this R24 grant program in March 2005, announcing a family of grants. Among this group, the Testing and Evaluation grant, a form of pilot-project deployment, is considered a research grant and funded through PHS 301. Six applications were reviewed in FY 2005, and one was funded. The award score was 154.

Resource Grants (MLAA)

Resource Grants, authorized by the Medical Library Assistance Act, support access to information, connect computer and communications systems, and promote collaboration in networking, integrating, and managing health-related information. Three of the four Resource Grant programs are centered on optimizing the management of health-related information; they are not research grants and are reviewed with relevant criteria. The fourth program, grants for Scholarly Works, supports the preparation of scholarly manuscripts in health sciences and health public policy areas.

Internet Access to Digital Libraries Grants

This grant program expired after the initial RFA in 2002 but was left open informally through FY 2004. Because of overlap with other EP grant programs and with activities of the NN/LM, this program was deactivated in FY 2005.

Knowledge Management and Applied Informatics Grants: In November 2004 a new program announcement was issued to replace the old Information Systems Grant program, which had expired technically and conceptually. The new program expands earlier goals of this resource grant program to emphasize projects that "translate" informatics research into practice. Fifty-seven grants were reviewed in FY 2005, compared with 84 in FY 2004. Seven awards were made. One was deferred for funding in FY 2006. The average score of awards was 145.

Integrated Advanced Information Management Systems Planning Grant

In keeping with NIH policy to update program announcements every three years, in March 2005 a new program announcement was issued for IAIMS Planning Grants. Minor modifications, including an explicit outline of the expected deliverable, were made to the program text. Twelve planning grants were reviewed in FY 2005, compared to 14 in FY 2004. One was funded and two were deferred to FY 2006 for funding. The average score of awarded grants was 163.

Integrated Advanced Information Management Systems Operations Grants

In keeping with NIH policy to update program announcements every three years, a new program announcement was issued for IAIMS Operations Grants in March 2005. Due to intense national interest in the installation of electronic health records, many calls were fielded, but no viable applications were received in FY 2005. Four applications were rejected as out of scope or unresponsive to the program. An IAIMS proposal can include clinical information systems within the information management system proposed by the project, but EP will not provide funds for the purchase and installation of such systems.

Grants for Scholarly Works

In November 2004 a new program announcement was issued to replace the expired Publication Grant Program. The new announcement clarified the scope and priorities of the program. Fifty-two applications were reviewed in FY 2005 compared to 70 in FY 2004. Ten

new awards were made, and five were deferred for payment in FY 2006. The average priority score of new publication grants funded was 142, compared to 167 in FY 2004.

Training and Fellowships (MLAA)

Overview

Exploiting the potential of information technology to augment health care, biomedical research, and education requires investigators who understand biomedicine as well as fundamental problems of knowledge representation, decision support, and human-computer interface. NLM remains the principal source of support nationally for research training in the fields of biomedical informatics. EP provides both institutional and individual training support.

EP-Supported Training Programs

Five-year institutional training grants support pre-doctoral, post-doctoral, and short-term trainees in 18 programs across the country. Seven of these programs were funded for the first time in 2002, while 11 are continuations of previously funded programs. While 15 programs involve a single institution, three are consortial programs that support training at multiple institutions. In total, 27 institutions provide training with NLM funds. Approximately 275 full-time trainees and 15 short-term trainees are being supported. Collectively, the programs emphasize training in clinical informatics, bioinformatics and computational biology, and public health informatics. EP receives some co-funding from the National Institute of Biomedical Imaging and Bioengineering (NIBIB) to support training related to advanced imaging methods at UCLA, and the National Institute of Dental and Craniofacial Research continues to help support training in dental informatics.

The 18 programs currently funded are at the following universities: California (Irvine), California (Los Angeles), Columbia, Harvard, Indiana, Johns Hopkins, Minnesota, Missouri, Oregon Health Science, Pittsburgh, South Carolina (Charleston), Stanford, Rice, Utah, Vanderbilt, Washington, Wisconsin, and Yale. The Harvard, Rice, and Minnesota programs involve multiple institutions.

Every summer, all NLM-supported trainees attend a national informatics training meeting. By tradition, this meeting takes place on the NIH campus in odd years, and at one of the funded training sites in even years. In 2005, the two-day meeting took place in the conference facilities of the NIH Clinical Center. It was attended by 350 trainees, program faculty, and NLM personnel. The meeting featured trainee research

presented in podium and poster sessions. It also included presentations of research under way at NLM, tours of the library, and open-houses hosted by the Lister Hill Center and the NCBI.

This program will be re-competed in FY 2006. Sites funded in next cycle will be supported from 2007 to 2012. To provide EP with a timely overview of what the programs are doing and thus help to structure the Request for Applications (RFA) for the re-competition, EP staff visited all 18 programs in 2003–2005. In advance of the visit, each site provided statistics and other background information on its curriculum, students and faculty. Following each visit, the principal investigator receives a letter which summarizes the findings of the team and becomes part of the official grant file. At this writing, the RFA for the re-competition is scheduled for release in January 2006.

In the past year, EP and the Robert Wood Johnson Foundation formed a partnership to lend increased emphasis to training in public health informatics. Through a $3.6 million grant from the Foundation to EP (through the Foundation for NIH), four existing training sites received supplemental awards to develop formal training tracks in public health informatics and for support of trainees in these tracks. The four sites were selected via an internal competition to which 11 sites applied. The four selected sites were Columbia, Johns Hopkins, Utah, and Washington.

Individual Fellowships

Informatics research training: EP offers two types of fellowship for informatics research training: an individual fellowship for basic or applied research (F37), which can be pre-or post-doctoral, and a senior fellowship intended for those with seven or more years of professional employment experience in an appropriate field (F38). In FY 2005, 19 applications were received for the F37 program, of which four were awarded. Nine applications were reviewed for the F38 program, and two awards were made.

Training for Informationists: In October 2003, EP issued program announcements for two new fellowships to support the training of in-context information specialists. These programs use the F37 and F38 mechanisms, but emphasize training for information specialist professional careers as opposed to research careers. Three F37 applications were received and three were funded. Three F38 applications were received and one was funded.

IAIMS Fellowships: The IAIMS fellowship expired in FY 2004 and was not renewed. Since its issuance, updates to EP's F37 and F38 fellowship programs made this fellowship unnecessary.

Career Support (MLAA)

Early Career Development Awards: The K22 program was established to provide transition assistance with funds for salary and for research to biomedical informaticians who are establishing their initial independent research programs. Twenty-two applications were reviewed in FY 2005, compared to 14 in FY 2004. Eight awards were made. The average score of a successful application was 157, compared to 165 in FY 2004.

Loan Repayment Program: EP participates in the NIH loan repayment program by identifying applications from informaticians involved in research related to clinical medicine. These applications are reviewed for merit by a Special Emphasis panel. For FY 2005, EP funded five of 11 applications.

Pan-NIH Projects

EP and Roadmap Activities

A major pan-NIH enterprise initiated by the NIH Director is resulting in a battery of RFAs and RFPs related to three themes: New Pathways to Discovery, Research Teams of the Future, and Reengineering Clinical Research. EP is a participant in all of these Roadmap initiatives. EP staff were particularly involved in NIH Roadmap teams for the National Centers for Biomedical Computing (NCBC) and in a number of interdisciplinary research initiatives.

NCBC and BISTI

The National Centers for Biomedical Computing are funded under renewable cooperative agreements. Although NIH Roadmap grants are considered pan-NIH grants, and awards will be managed by teams from a number of Institutes, each Roadmap grant has a "home" Institute, which provides the program officer for the Center. An initial competition resulted in the funding of four NCBCs in September 2004. One of these centers, "Informatics Integrating the Bench and Bedside," based at Harvard's Brigham and Women's Hospital, is administered through EP. In FY 2005, a second round of competition resulted in funding three additional centers. An NLM intramural scientist is the lead science officer for one of the new NCBCs, based at the University of Michigan.

The NCBCs are intended to operate as a network, creating a national infrastructure for advanced biomedical computing. To this end, the centers are administered through a project team in which several EP professional staff participate. Like all NIH Roadmap initiatives, the NCBC program will undergo formal and

systematic evaluation study. An EP staff member has taken the lead for the pan-NIH project team in planning the evaluation of the NCBC effort. Although conceptually related to the NIH BISTI (Biomedical Information Science and Technology Initiative) program, the NCBC program is a distinct program, funded through NIH Roadmap grants.

Multi-institute Program Announcements

In addition to its involvement in the NIH Roadmap, EP also participates with other NIH and federal organizations in a number of multi-agency projects, including several types of BISTI grants. The applications for these programs are reviewed by CSR, and then participating institutes select grants for full or shared funding. EP participation has been steady but is rarely more than one new grant each year, and in some years none is funded. The statistics for these programs are folded into regular grant program counts; most are R01 grants. An updated listing of the multi-institute initiatives in which EP participates is available on the EP Web site at http://www.nlm.nih.gov/ep/ GrantsOther.html.

Shared Funding for Research & Training

EP contributed approximately $1 million in collaborative co-funding agreements to grants in FY 2005. The listing below outlines the projects and amounts for research co-funded with other NIH institutes and centers:
- Co-funding ($200,000) for the Protein Databank at Rutgers University. This databank supplies 3-dimensional representations of proteins and is the single worldwide archive of structural data for biological macromolecules.
- Co-funding ($300,000) with the National Institute of General Medical Sciences for continued support of a cooperative agreement for the Stanford Phamacogenetic Research Network and Knowledge Base.
- Co-funding ($25,000) with the National Institute of Neurological Disorders and Stroke (NINDS) for a project titled, "A Mature Brain Architecture Knowledge Management System." The objective of this project is to develop a user-friendly neuroinformatics workbench for the Web allowing the neuroscience community to access, evaluate and visualize neuroanatomical literature.
- Co-funding ($173,000) to the National Center for Research Resources (NCRR) in support of a Neuroimaging Analysis Center that would support further development and extensions of the Insight Toolkit (ITK) that will be used in working in a grid computing environment. This grid computing

application will allow for computation of very large datasets that otherwise could not be analyzed effectively.
- Co-funding support to NCRR for the pan-NIH initiative on Electronic Research Administration (eRA). This cooperative agreement, "Electronic Submission of Grant Applications," supports the overall Federal Government's e-Grants and e-Government initiatives.
- Co-funding ($250,000) for two grants in the Fogarty International Center's Informatics Training for Global Health, administered by the FIC.
- Co-funding ($23,132) with the National Institute on Aging for a grant entitled, "Neuroprotective Apolipoprotein-E Analogs." The goal of this project is to develop neuroprotective treatments to limit and prevent cognitive and behavioral impairments in patients with traumatic brain injury.
- Co-funding ($52,130) with National Human Genome Research Institute (NHGRI) for an STTR grant titled "Development of a Portable PME DNA sequencer."
- Co-funding ($25,870) with the National Heart Lung and Blood Institute (NHLBI) for an STTR proposal entitled, "Cord Blood: Treatment for Acute Myocardial Infarction".
- Co-funding ($100,000) with NHGRI for a research grant entitled, "BioMediator: Biologic Data Integration & Analysis System" for searching across various genomic databases for the purpose of curating the GeneClinics Database formerly supported by a NLM grant. GeneClinics has now been integrated into the NCBI as a publicly available resource for clinicians and scientists.
- Co-funding ($63,000) with NHGRI for a grant entitled the "Oral History of Human Genetics: The Intelligent Archive" that will include a collection of over 100 oral histories from clinicians, scientists, theorists, organizational leaders and others covering the ethical, legal and social issues surrounding the field of human genetics.

Interagency Agreements (Specialized Agreements)

EP partnered with the National Science Foundation (NSF) on a new grant program titled "Dynamic Data-Driven Systems." EP provided funding for one grant in this program: "Dynamic Data-driven Brain Machine" ($336,877). The project's goal is to develop computer architecture for computational modeling and training algorithms for robotic arm movement. A clinical application of this basic work relates to improving the brain-machine interface for paraplegics. For this

partnership, the grant is administered by NSF, with EP providing adjunctive program officer oversight.

In the multi-agency "Multi-scale Modeling in Biomedical, Biological and Behavioral Systems," EP partnered with NSF, National Aeronautics and Space Administration, Department of Energy, and eight NIH Institutes. In addition to sharing costs for the grant review at NSF, EP funded one grant ($364,755) in this program, which was transferred from NSF to NIH. This grant, titled "A multi-scale approach for understanding antigen presentation in immunity," is administered by EP as an R01 grant.

SBIR/STTR (PHS 301)

All NIH research grant programs allocate a fixed percentage of available funds every year to Small Business Innovation Research (SBIR) grants. These projects may involve a Phase I grant for product design, as well as a Phase II grant for testing and prototyping. SBIR and STTR applications are reviewed by CSR. Fifty SBIR/STTR applications (49 of them Phase I applications) were assigned to EP in FY 2005 and reviewed by CSR. Six awards were made. Of these applications, 25 were "unscored," indicating reviewer assessment that they were not competitive for funding.

EP Operating Units

Committee Management Office

The Board of Regents approved 166 grant applications. The Board Operating Procedures were reviewed and approved without change at the February 15–16, 2005 meeting. Extramural Programs staff made a number of presentations to the Board of Regents in FY 2005. A summary of the Board's activities is in the "Board of Regents" section of the chapter on Administration.

Grants Management Office

The NLM, with a reduced grants management staff, issued 221 grant awards in FY 2005 for over $53 million. Staffing levels will be restored in FY 2006. NLM provided key support for all grant co-funding agreements of the NLM, interagency agreements in support of grants, large scale training grants, and its general training and research grants split over two funding authorizations, the Medical Library Assistance Act and the PHS Act.

Policy required that the grant from the Robert Wood Johnson Foundation for support of public health informatics, be made to the NIH Foundation. The Grants Management Office took the lead in developing the negotiated agreement. The awards are handled as

supplements to four of the 18 approved training programs.

In addition, the NLM Grants Management Office provides full grants accounting support for the EP budget for all awarding mechanisms, updates the annual Catalog of Federal Domestic Assistance, and works closely with the NLM Freedom of Information Coordinator to provide timely response to FOI requests for grant information.

Program Office

Program activities in FY 2005 were focused on updating EP's grant program announcements, evaluating the university-based informatics training programs, building collaborations, and publicizing EP's grant programs. EP's Web site was updated with awards listing after each payline. Also, program staff identified all current and expired programs and developed a timetable for re-issuing expired programs. Revised program announcements were issued for five programs: IAIMS Planning, Operations, and Testing & Evaluation grants; Knowledge Management / Applied Informatics grants (formerly Information System grants); and Grants for Scholarly Works (formerly Publication grants).

Working with DEAS staff, EP implemented a new process of administrative review for incoming grant applications. Corrections are now made to Program Officer and Program Class Codes in IMPAC at the time EP accepts assignment of an application. A new notification program was instituted so that applicants would know their grant number and assignment to NLM.

EP Staff and Dr. Harold Schoolman (retired NLM Deputy Director for Research and Education and now a consultant to NLM), completed the cycle of evaluation visits to EP's 18 training programs. A panel proposal on the study was accepted by American Medical Informatics Association, to be presented at the fall AMIA meeting, October 24, 2005.

EP became the coordinating center for one of NIH's Government Performance and Results Act goals. The 3-year goal is titled CBBR-5: By 2007, Expand by 5,000 the Pool of Researchers & Clinicians NIH has Trained in Biomedical Informatics, Bioinformatics, or Computational Biology. This activity requires three reports each year during the life of the goal, drawing together information from NLM extramural and intramural programs, NIBIB, NIGMS and NHGRI.

Increasingly, EP's research grants involve human subjects. The grants which are clinical trials require population tracking. This activity has not previously been undertaken at EP, but will begin in FY 2006. In FY 2005, planning began for deciding which grants would be tracked and how the data would be entered into IMPAC.

Scientific Review Office

EP's initial review group, the Biomedical Library and Informatics Review Committee, evaluates grant applications assigned to EP for possible funding for scientific merit. BLIRC met three times in FY 2005 and reviewed 177 applications. The Committee (Appendix 5) reviews applications for most biomedical informatics and bioinformatics research applications, knowledge management/applied informatics, career support and fellowships. BLIRC has two standing subcommittees: the Networked Information Access Subcommittee, and the Medical Informatics Subcommittee. The subcommittees review fellowship applications: informationist in the former committee, informatics in the latter. The charter of the BLIRC was amended to reflect the broader scope of research applications in the areas of clinical informatics, bioinformatics, biomedical computing, management of health science information, as well as library science.

Special Emphasis Panels: Twenty-two Special Emphasis Panels were held during FY 2005. These panels are convened on a one-time basis to review applications for which the regularly constituted review group lacks appropriate expertise, or when a conflict of interest exists between the applicant and a member of the BLIRC, or for certain special categories, such as Conference grants. The Scholarly Works grants are routinely reviewed by a panel, due to the unique expertise needed. These panels reviewed a total of 248 applications during FY 2005.

A second level peer review of applications is performed by the Board of Regents. One of the Board's subcommittees, the Extramural Programs Subcommittee, meets the day before the full Board for the review of "special" grant applications. Examples include applications for which the recommended amount of financial support is larger than some predetermined amount; when at least two members of the initial review group dissented from the majority; when a policy issue is identified; and when an application is from a foreign institution. The Extramural Programs Subcommittee makes recommendations to the full Board, which votes on the applications. The Board also votes en bloc for all other applications that meet criteria for further consideration for funding.

Table 12

EXTRAMURAL PROGRAM BUDGET

	NON COMPETING		COMPETING		TOTAL	
	NO	AMT	NO	AMT	NO	AMT
MLAA		($)		($)		($)
IAIMS (G08)	8	$2,101	2	$239	10	$2,340
TRAINING						
TRAINING PROGRAMS (T15)	18	$14,041	0	$0	18	$14,041
FELLOWSHIP/OTHER (F37/F38/D43)	8	$558	11	$730	19	$1,288
CAREER(K22)	5	$777	8	$1,220	13	$1,997
TOTAL TRAINING	26	$15,376	19	$1,950	50	$17,326
SCHOLARLY WORKS (G13)	21	$1,376	9	$636	30	$2,012
RESOURCE						
IADL(G07)	2	$58	4	$322	6	$380
INFO. SYS.(G08)	29	$3,691	5	$691	34	$4,382
TOTAL RESOURCE	31	$3,749	9	$1,013	40	$4,762
LOAN REPAYMENT (L30)	0	$0	5	$365	5	$365
NN/LM CONTRACTS (N01)	8	$12,442	0	$0	8	$12,442
TOTAL MLAA:	94	$35,044	44	$4,203	138	$39,247
PHS 301						
BIOMED-INFORM. RESEARCH (R01/R03/R13/R21/R24)	28	$6,851	22	$5,254	50	$12,105
INTERAGENCY AGREEMENTS	1	$200	1	$337	2	$537
CHAIRMAN'S GRANT(U09)	1	$87	0	$0	1	$87
BIOMED-INFORM. RESEARCH TOTAL	30	$7,138	23	$5,591	53	$12,729
BIOINFORM. RESEARCH(R01/R03/R21)	14	$5,060	4	$1,088	18	$6,148
BIOINFORM. RESOURCE *commited* (P41)	5	$3,133	0	$0	5	$3,133
BISTI(R01/R21/R33/U54)	7	$3,600	5	$1,034	12	$4,634
BIOINFORM. RESEARCH TOTAL	26	$11,793	9	$2,122	35	$13,915
SBIR/STTR(R43, R41, R42)	0	$0	6	$728	6	$728
TOTAL PHS 301:	56	$18,931	38	$8,441	94	$27,372
TOTAL EP:	150	$53,975	82	$12,644	232	$66,619

Budget excludes NIH Taps

OFFICE OF COMPUTER AND COMMUNICATIONS SYSTEMS

Simon Y. Liu, Ph.D.
Director

The Office of Computer and Communications Systems (OCCS) provides efficient, cost-effective computing and networking services, application development, technical advice, and collaboration in informational sciences to support NLM's research and management programs. OCCS develops and provides the NLM backbone computer networking facilities, and assists other NLM components in local area networking. The Division provides professional programming services and computational and data processing to meet NLM program needs; operates and maintains the NLM Computer Centers; develops software; and provides extensive customer support, training courses, and documentation for computer and network users. OCCS helps to coordinate, integrate, and standardize the vast array of computer services available throughout all of the organizations comprising NLM. The Division also serves as a technological resource for other parts of the NLM and for other Federal organizations with biomedical, statistical, and administrative computing needs.

Business Continuity and Disaster Recovery

In order to protect NLM's mission-critical systems, CIT and NLM have implemented an NIH Consolidated Collocation Site (NCCS) in Sterling, Virginia. The NCCS is now in operation as a disaster recovery/alternate computing site for NLM as well as CIT, NCI, NHLBI, NIDDK and NIGMS. At present, all NLM mission-critical systems are either under active/active, active/passive or active/cold-backup mode depending on their business requirements. The Business Continuity and Disaster Recovery Plan covers NCCS as the primary resource for system restoration and uninterrupted processing if the primary NLM computing facilities on the NIH campus are rendered unavailable by a disaster or other contingency. During this year, OCCS deployed the DOCLINE application at the NCCS. This deployment consisted of the DOCLINE front-end Web and Cold Fusion processors operating in an active/active mode with the Oracle database active at NLM for now.

The NLM computer facility has tripled its use of electrical power, cooling and data transmission capacity over the last three years due to the dramatic growth in dependence on IT systems to deliver NLM's mission-critical applications. Recognizing that this rapid growth will continue in the years ahead, OCCS has begun a detailed reengineering process for evaluating the safety, reliability and performance requirements of the computer facility. The reengineering efforts include:

- Expanding the Uninterrupted Power Supply (UPS) capacity to support the growing needs for electrical power protection and redundancy of systems housed in the NLM computer facility. The computer facility currently can maintain electrical power for up to 39 minutes after losing commercial power.
- Developing plans for an overhead Ladder Rack in the computer facility, as a separate pathway for running data networking cables to improve the reliability, availability, and maintainability of data communication services.
- Initiating plans for the implementation of a pre-action sprinkler system to improve the reliability and safety of the fire suppressing system. The current system is a traditional wet pipe system. The proposed pre-action sprinkler system would require two actions before water will be released onto the fire: first, the smoke detection system must identify a developing fire and then open the pre-action valve; second, the sprinkler head must release to permit water onto the fire.
- Developing plans to re-engineer the computer facility electrical and grounding rail system. When completed, the electrical power will increase by 25 percent while decreasing the time of delivery of electrical power to the IT systems.

Consumer Health

MedlinePlus: MedlinePlus's greatest achievement this year was the enhancement of the MedlinePlus Go Local feature which brings local health services to the public by allowing users to search for healthcare providers in their localities while searching MedlinePlus for information on medical conditions or other healthcare issues. Originally, Go Local information was hosted only on local Web sites linked to MedlinePlus Go Local. This year, for the first time, OCCS fielded Go Local sites built and customized by the final version of the Go Local Input System. This allows localities lacking their own Web resources to enter/maintain data and select look-and-feel elements for Go Local pages that NLM hosts for them.

MedlinePlus added Go Local sites serving Indiana, Massachusetts, Southern Texas, and Alabama. Existing Go Local sites are North Carolina and Missouri. Both the Indiana and Massachusetts sites were generated by the Go Local Input System, a true leap forward in the quality and level of MedlinePlus service. Additionally this year, MedlinePlus added: health check

tools in both English and Spanish; organizations pages grouped by health topic; a search form embedded in all consumer mailing list messages that permits email recipients to search MedlinePlus directly from their inboxes; and A–Z shortcut links for in-page navigation in email. MedlinePlus realized a 35 percent increase in page views with 645 million pages views and 73 million unique visitors.

SeniorHealth Project: SeniorHealth is a joint NLM and National Institute on Aging (NIA) project that provides health information on the Web using modes of delivery video and narration appropriate for older Americans with low vision and/or low hearing. The system includes the Accent "Talking Web" module developed by OCCS to provide accessibility enhancements, including a selectable range of type sizes and spoken text.

Several new topics were added to SeniorHealth, including Diabetes, Taking Medicines, and Exercise. Baseline access was broadened to comprise additional NIH Institutes, including the NEI, NINDS, and NIDCD. Added for these Institutes were information on age-related eye problems, problems with taste and smell, and strokes. The Accent module received numerous enhancements, including improvements to server security and the ability to have spoken text begin as soon as a page is loaded. SeniorHealth received two awards, from the International Council on Active Aging as an Industry Innovator, and an Honorable Mention from Medicine on the Net in the category of Best Overall Consumer Portal.

Local Legends: Local Legends, the companion Web site to the Changing the Face of Medicine exhibit, features biographies and video clips of outstanding American women physicians who have been nominated by members of Congress. This year, OCCS conducted several rounds of Local Legends usability testing and made numerous modifications to improve usability, including improvements to navigation menu options, development of more succinct text for Local Legends pages, implementation of an organized women physicians' overview page, and many other clarifications and enhancements. A guestbook and a registration process were also added.

Images from the History of Medicine (IHM): More than 500 images were added to the IHM database. IHM's currently installed backend software was replaced with a product from Luna. MARC output has been exported from the IHM database and quality assurance reports have been produced.

Daily Med Project: The DailyMed project is a partnership between the Food and Drug Administration, the Veterans Administration, the NLM, medication manufacturers and distributors, and healthcare information suppliers. The project seeks to provide a standard, comprehensive, up-to-date, XML-based capability for labeling the contents of medications. This year, OCCS:

- Developed mockup pages for DailyMed Web application interfaces;
- Designed DailyMed database, transaction catalogs, and user registration database schemas;
- Designed and developed processing programs at NLM, including daily-upload verification, data processing, and error notification algorithms; and
- Defined and implemented DailyMed system redundancy architecture.

IT Security

NLM continued to assess and strengthen its security posture based on current business requirements and risk assessment. Security improvements continued throughout the year. OCCS continues to perform a monthly cycle of vulnerability scanning, detection, and remediation thereby making concrete improvements in NLM's security posture.

NLM implemented a new NIH password policy. The policy and password requirements enforce password complexity, password change frequency, and issues related to password caching, sharing, and system inactivity locking. Guidelines were distributed to staff and user awareness training was conducted.

In response to a request from the NIH Information Security Awareness Office, NLM reviewed and revised the Security, Risk Assessment, and Self-Assessment Plans for mission-critical systems as well as the Business Continuity and Disaster Recovery Plan for NLM as a whole. The Plan has been modified to include procedures for recovering designated NLM medical computing systems including resumption of activities at the NCCS if the primary computing facilities on the NIH Campus become unavailable.

NLM implemented a new NIH policy governing the use of the Simple Mail Transfer Protocol Server. This policy specifies that incoming and outgoing traffic on this server will be scanned by Central E-mail Services or by the NIH mail forwarding service. Only these servers will be authorized to distribute or receive e-mail across the NIHnet perimeter firewalls. NLM has blocked all unauthorized traffic at our firewalls and complies with the policy. All government agencies must support strong authentication using Public Key Infrastructure. Two NLM Security staffers in OCCS became Local Registration Authorities and provided certificates for use with email and encryption systems.

In October 2004, the evaluation of Symantec's antivirus offering was completed. The evaluation produced a very positive outcome resulting in the replacement of McAfee Anti-Virus with Symantec Anti-Virus as NLM's advanced, enterprise-wide virus protection.

The Office of Management and Budget requires that HHS computer users complete annual IT security awareness training. NLM has completed 100 percent of the mandatory FY 2005 Security Awareness Training for employees and contractors.

Professional Health Information

RxNorm Project: Two versions of the RxNorm Editing System were released in FY 2005. Release 2.5 included full integration of the RxNorm user interface design framework and a major reworking of the mechanism for creating the "normalized form," or standardized form, of the generic drug name. Major enhancements to the Editing System in Release 3 included the ability to search for and choose certain types of forms from a list and the enabling of hot keys on the Component Editing pages. A total of four releases of the RxNorm nomenclature itself (as opposed to the Editing System) occurred in this year. RxNorm is resynchronized with UMLS data with each UMLS release.

Unified Medical Language System (UMLS) Project: The Unified Medical Language System (UMLS) provides a common, concept-oriented medical vocabulary and thesaurus based on more than 100 current medical source vocabularies in a variety of languages. UMLS has been a major project of the National Library of Medicine for more than a decade. During FY 2005, NLM began transferring support for the UMLS Metathesaurus from the Lister Hill Center to OCCS. A production environment for the system was established and the necessary program modules were tested and put in place. The program code was refactored to streamline its design and make it function well in the OCCS environment. By the end of the fiscal year, Phase I production was being carried out in shadow mode.

Medical Subject Headings and Related Systems: The M2000 client-server application used to maintain the Medical Subject Headings database received a number of significant enhancements, including a new Qualifier Maintenance Module, enhancements to the Post-authorization Module, and a new Archive System. This year, new MeSH Translation Management System (MTMS) modules were created for distribution of MeSH data in other languages, including French, Italian, German, and Japanese. Additionally, MTMS received a major upgrade that allows a translation to synchronize with previous years of English MeSH.

Data Creation and Maintenance System: The major event this year for the Data Creation and Maintenance System (DCMS) is the baseline extract, a re-release of all DCMS citations with the new MeSH headings. A new Java-based Loader/Extractor module now loads data into DCMS from XML files, maps the XML elements onto the structure and contents of the DCMS database, and extracts DCMS data to output XML files. With the help of the new Loader/Extractor, the baseline extract was completed this year in a record 32 hours. OCCS also:

- Added support for the ClinicalTrials.gov Web site;
- Removed keyboarding as a data creation method from the DCMS interface (the system retains support for keyboarding in case this is needed in the future);
- Upgraded system software to support the new MEDLINE DTDs—the XML Data Type Descriptions that provide standard data import and export for the MEDLINE database;
- Provided workflow for OLDMEDLINE citations—citations dating from 1950 through 1965—which are now included in DCMS;
- New code to bypass MeSH validation for OLDMEDLINE records (which preceded MeSH); and
- Added Meeting Abstract legacy data and modified the New Article function to support OLDMEDLINE and Meeting Abstracts.

DOCLINE: DOCLINE, the NLM interlibrary loan (ILL) system, supports approximately 3,600 domestic and international libraries in processing more than 2.5 million interlibrary loan transactions a year. After exhaustive testing by two university libraries, DOCLINE 2.3 was released early in FY 2005 providing the capability of using the ILL protocol from the International Standards Organization to communicate data between proprietary ILL systems and DOCLINE. Two more DOCLINE 2.3.x versions were released this year. A maintenance release (2.3.1) was followed by Version 2.3.2, featuring the DOCLINE Data Warehouse. Released in June 2005, Version 2.5 of DOCLINE provided a complete redesign of the Loansome Doc ordering system which allows users to order full-text copies of articles from a DOCLINE Network medical library. Over 40 DOCLINE enhancements were implemented during the fiscal year.

Public Health Partners (PHPartners.org): Public Health Partners is a collaboration of U.S. government agencies, public health organizations and health sciences libraries to present information for the public

health workforce on a single Web site. Version 2 of the PHPartners was implemented this year. A tutorial entitled "Public Health Information and Data" was created for PHPartners by the University of Michigan. A new Public Health Knowledge page was also brought online.

Health Services Research Information (HSRInfo) Central: FY 2005 saw the release of a new Health Services Research (HSR) Portal: Health Services Research Central, or HSRInfo Central. HSRInfo Central provides links to Web sites of interest to the users and producers of health services research. The links are organized into several categories, including Data Tools and Statistics, Grants and Funding, and News. The site made its debut in June 2005 and has generated enthusiastic feedback. The PHPartners and HSR Portal Input System, a new system for maintaining both PHPartners and the HSR Info Central database was implemented this year.

Health Services Research Projects in Progress: A new version of the Web interface to the National Information Center on Health Services Research and Health Care Technology (NICHSR) Health Services Research Projects in Progress database was implemented this year. The new interface includes greatly expanded search capabilities and an improved Help function.

Voyager Integrated Library System (ILS): A validation system for Voyager was completed and made available both from the internal NLM network and through Citrix. A new release of Voyager bibliographic data in XML was completed and the full set of over 700,000 records was provided to NCBI. NCBI loaded the records for searching by Entrez. The specifications are now being modified and a final version of the input system is expected to be available early in FY 2006.

Scan Track (PubMed Central Inventory): PubMed Central (PMC) is NIH's free digital archive of biomedical and life sciences journal literature. In FY 2005, NLM supported PMC with a new service called Scan Track, which uses an Oracle database to track the status of articles through the indexing process and inclusion in PMC.

Literature Selection Technical Review Committee (LSTRC): In FY 2005, a new MEDLINE Review application was implemented. LSTRC uses this new application to review journals for inclusion in MEDLINE.

Serials Extract File (SEF): Among numerous upgrades and fixes, the loading of Serials XML data for NCBI was completed and OCCS added a number of new fields to the SEF database. A LSTRC-to-Serial Subset Membership process was implemented. The NCBI journals database was updated using a modified DTD.

Network and Systems Support

OCCS continued to provide reliable LAN and Internet communications services, meeting the data communications needs for new IT systems, providing security services as well as end user assistance and training, implementing new network-based applications and operating systems, and exploring new technologies and plans to meet NLM's continued growth in networking and communications. OCCS took steps in a number of areas to increase the capabilities and reliability of network services:

- NCCS data communications services
- Enhanced network monitoring and management
- Increased IT and network security
- New networked services to support the NLM user community
- Additional redundancy to eliminate single points of failure
- Enhanced backup for use in disaster recovery scenarios
- Expanded and efficient centralized shared data storage

Public Internet connectivity services to NLM are provided through a contract with Level3. Internet connectivity is provided via an OC3 (155 Mbps) circuit to the Level3 network node in McLean, VA. The contract also provides an OC3 link for CIT/NIH to the Level3 network. NLM and NIH use these connections to back up each other's Internet connectivity. The service features an automatic failover in the event of a scheduled or unscheduled outage of one Internet connection.

A remote access Citrix terminal server solution continues to be an effective solution for NLM flexi-place workers, as well as staff and contractors who need temporary or long term remote access to NLM IT systems and applications. The terminal server system provides authentication into the NLM network, access to office and NLM business applications, network-based files, and the Internet. High-speed access is provided mainly through cable modems provided by COMCAST.

Wireless services were implemented throughout key areas of Buildings 38 and 38A. Wireless access to the Internet and public services of NLM and NIH is provided for guests and typical users. Through a Virtual Private Network, authorized users can access internal applications in a secure manner.

Internet2 has become an important resource for connection with NLM and the research community. Internet2 connectivity is provided by a Gigabit Ethernet

link to the Abilene high-speed backbone network via the Mid Atlantic Exchange at the University of Maryland. LHC and OCCS cooperate to manage traffic to and from Internet2. There are plans to implement a redundant, diverse connection via Fibergate dark fiber to provide increased reliability for this critical connection.

This year, OCCS implemented the High Availability Computing Solution to ensure that critical applications and resources remain available to NLM users. OCCS deployed clustered Oracle server systems and clustered storage systems as NLM's high availability computing resources. A server cluster is a group of independent computer systems working together as a single system thereby allowing multiple servers to deliver the same application services so that if one of the servers becomes unavailable another immediately begins providing service. This initiative will increase the NLM mission critical database storage capacity eight times and significantly reduce the backup cycle time via clustered network attached storage technology.

OCCS continued to make improvements to the UNIX architecture. Various upgrades in additional servers, increased memory, and subnet reliability were performed.

Desktop Support

OCCS once again championed NLM's participation in a licensing agreement that provides a bundle of Microsoft products at the lowest cost available in the U.S. The original Maryland Education Enterprise Consortium contract, which NLM joined in 2001, expired in June of this year. The NLM Personal Computer Advisory Committee approved a recommendation to adopt a new agreement, and OCCS assisted with the negotiation of the contract vehicle. NLM has now joined a new three-year, two-option-year term agreement for acquiring Microsoft software and operating system licenses and updates. Under the new package agreement, seats are priced at $18.50 per FTE. By contrast, GSA prices for the same package total over $1,712. With current licensing needs extended over the possible 5-year life of the agreement, this new vehicle is expected to save over $2 million dollars in licensing and maintenance fees.

PestPatrol, a centrally managed spyware/adware detector and remover, was deployed across OCCS-supported systems. This product automatically removes, on a weekly basis, thousands of instances of nuisance or malicious desktop objects. Reporting to a central console permits OCCS to monitor and correlate this activity to track risks to further infection of NLM systems.

To adopt the latest available e-mail client, Outlook 2003, OCCS developed and deployed the upgrade to the entire MS-Office XP suite, which includes Word, Excel, PowerPoint, and MS-Outlook. This deployment was scheduled office-by-office to ensure that any impacts could be quickly resolved. NLM staff received training to ensure an effective migration.

Four hundred and twenty security patches are now consistently applied to OCCS-supported desktops running the Windows operating system. The Software Update Server security hotfix deployment solution introduced by OCCS makes possible the expeditious deployment of critical security updates, keeping 1,200 NLM systems better insulated from attack.

Research and Development Initiatives

Voice Recognition: Two-way voice communication is an emerging technology that enables users to navigate Web sites by means of spoken user input and interact with Web applications that synthesize speech from the Web server. This year, OCCS investigated grammars and signal repositories for the Apple, Windows, Linux, and UNIX platforms. As this technology develops, it promises to add a vital new resource for convenience and accessibility to the NLM Web.

Database: OCCS initiated the Oracle High Availability Project, which aims to move OCCS systems to Oracle Database 10g with Real Application Cluster. This is a cluster storage management system at the application level. It uses a shared-disk architecture that allows applications to grow in scale as needed. Existing applications can use the database without any modifications.

ReportNet Migration: An order agreement and plan was developed for switching from Impromptu and Impromptu Web Reports to the newer Cognos product ReportNet. The Cognos reporting products are used for numerous decision-support reporting tasks.

NLM Web Support

Web Content Management: NLM uses TeamSite to provide content and application management for Web sites. There are plans to upgrade to Version 6.5. Also, a version control server was set up using the Subversion system. This provides version control for applications that do not fit the TeamSite model well.

Web Analytics: NLM uses WebTrends software to track the number of pages served over time by the sites being managed and to analyze trends in site usage, audience composition, and other matters. To improve performance, WebTrends 6.x was upgraded to WebTrends 7.5 Distributed Architecture.

Advanced Search Engine: OCCS devotes significant effort to maintaining and improving search capabilities for the NLM Web sites. This year, OCCS put in place a new Search Engine for Technical Bulletins, implemented a MedlinePlus Go Local and Servlets Interface for NLM's "one-stop shop" Gateway Version 2, updated the look and feel of the Local Legends Search Engine, and initiated a research project to perform a comparative study of available search engines.

Link Checker: Maintaining the validity of Web links is an important challenge for Web administrators. Using locally developed Java software that generates HTML reports from XML-like markup language produced by the system, Link Checker was implemented for Voyager and two Outreach applications.

Technical Bulletin: A small project, called Gold Standard, was implemented to receive and respond to users' questions about Technical Bulletins. Additionally, the Technical Bulletin templates were re-worked to be compliant with the Extensible Hypertext Markup Language format, resulting in improved accessibility and higher-quality PDF file production from the ColdFusion document generator.

Customer Services

Since the 2003 Help Desk consolidation with NIH's IT Help Desk, NLM desktop and PC networking support requests are now channeled to the NIH IT Help Desk for initial ticket entry into the call tracking system. This year over 9,600 NLM ticket requests for IT support were entered and tracked. NLM IT Staff resolved over 72 percent of the calls with 28 percent of support calls being completed by NIH staff.

OCCS conducted over 70 courses this year, in topics such as "SPAMology," "Outlook 2003 Overview," "Power-up PowerPoint," "Managing your Mailbox," and "Office 2003 Differences." Public briefings were also conducted in support of the MS-Office 2003 migration, Stay-In-School programs and NLM Associates programs.

Process Improvement Initiative

OCCS continued implementation of repeatable standards for all stages of the software life cycle. Standard operating procedures for daily practices were created on topics such as communications and network services, desktop systems support, systems administration, security and many other activities. In addition, process improvement audits developed 14

recommendations to improve efficiency and consistency.

The Desktop Services Section, participating with the NLM Personal Computer Advisory Committee (PCA), last year developed technical standards and product selections for two classes of notebook systems and a secure thin-client PC to join the PC Desktop selection in PCA consolidated orders. Within the range of laptop computing requirements, a single technical solution is not available to meet all portable PC needs. The PCA defined a third technical requirement for an "Ultralight" configuration, and an appropriate specification and configuration were developed. This joins the current specifications and configurations for "Powered-On" and "Mobile" notebook models.

OCCS continues to convene a Configuration Control Board (CCB) to provide oversight of configuration changes made to production IT systems managed by the Systems Technology Branch. Quality management is a top priority of OCCS and quality management improvements are expected to lead to significantly greater maturity and repeatability in the day-to-day operations of the Division.

Data Center Operations and Administrative Support

NLM systems continue to be supported in a safe environment in NLM's computer facility, which is available 24x7x365. The Network Operations and Security Center (NOSC), established in 2002, continues to serve as a central point in IT system and service monitoring, IT system administration, IT security event monitoring, and after-hours Help Desk support.

The NOSC display system consists of four 32-inch wide-screen plasma displays that are visible outside the computer room. The intended audience of this display system is the general public and NLM staff. The system consists of information "panels" with descriptive text, statistical charts and near real-time activity monitors. Each panel focuses on a particular NLM service or IT infrastructure component. The panels include near-real-time utilization counters for MedlinePlus and for PubMed/MedlinePlus, and NLM services as seen by remote users around the world. Near real-time utilization data for NLM's Internet1 and Internet2 data communications links are also displayed.

Customer Service Support System: Among numerous upgrades and enhancements, OCCS migrated the Customer Service Support System, Siebel, from Oracle 8i to 9i Unicode, making the database compliant with NLM standards. Additionally, the team completed work on a new Siebel XML integration capability for inbound Web form-based requests from the public.

ADMINISTRATION

Todd D. Danielson
Executive Officer

Table 13

Financial Resources and Allocations, FY 2005 (Dollars in Thousands)

Budget Allocation:

Extramural Programs	$69,400
Intramural Programs	240,971
Library Operations	(92,953)
Lister Hill National Center for Biomedical Communications	(61,258)
National Ctr. for Biotechnology Information	(72,227)
Toxicology Information	(14,533)
Research Management and Support	10,809
Total Appropriation	321,180
Plus: Reimbursements	10,251
Total Resources	**$331,431**

Personnel

In September 2004, **Michael Muin, MD**, joined the LHNCBC as a visiting scholar. Dr. Muin received his medical training at the University of the Philippines where he is a Professor of Clinical Anatomy. At NLM, he will be developing the Virtual Evidence Cart, an online tool that facilitates the practice of evidence-based medicine.

In September 2004, **Artem Novozhilov, Ph.D**., joined the staff of the NCBI Computational Biology Branch as a Visiting Fellow (VP). Dr. Novozhilov obtained his Master's in applied mathematics and his Ph.D. in physics and mathematics from Moscow State University. His expertise is in biologically relevant mathematical modeling. At NCBI, Dr. Novozhilov will do research on genome evolution, the goal of which is to reach a better understanding of which aspects of biological complexity embodied in network topologies are due to physical self-organization principles and which are the products of natural selection.

In September 2004, **Alissa M. Resch, Ph.D**., joined the staff of the Computational Biology Branch, NCBI as an IRTA fellow. Dr. Resch received her Ph.D. in biochemistry from the University of California, Los Angeles, in 2004. While at UCLA she researched the theory of molecular evolution, developing methods to ask a basic question about the role of intron insertion events in the evolution of protein domains (i.e. Pfam domains). At NCBI, Dr. Resch will perform research on evolution of eukaryotic gene structure, alternative splicing, and the emergence of eukaryotic genome complexity.

In September 2004, **Timothy Valin** joined the staff of the Office of Administrative Management, National Center of Biotechnology Information (NCBI) as an Administrative Officer (AO). Mr. Valin received his BS from the State University of New York at Albany (SUNY) and is currently completing his MBA at Mount St. Mary's University in Emmetsburg, MD. His background includes management of multiple private companies as well as government contracting. At NCBI, Mr. Valin will direct the operations of the Office of Administrative Management.

In October 2004, **Dianne McCutcheon** returned to NLM as chief of the Technical Services Division, Library Operations. Before this appointment, she was a systems planning analyst at the Library of Congress. She had earlier served NLM in a variety of positions including head of the Serial Records Section. She is a former NLM Associate and also has experience as a hospital library manager. In addition to her key role in NLM's ILS implementation, Ms McCutcheon was instrumental in the development of serials automation at NLM, including the national serials holding database.

In November 2004, **Evan Bolton, Ph.D**., joined the staff of the Computational Biology Branch, NCBI, as a Staff Scientist. Dr. Bolton received his doctoral degree in physical chemistry from the University of Georgia in 1995. Dr. Bolton worked as a computational chemist in the life science industry for over eight years. At NCBI, Dr. Bolton will focus his research efforts on developing the PubChem data system.

In November 2004, **Mehryar Ebrahimi** was appointed Chief of the Office of Administrative and Management Analysis Services. Mr. Ebrahimi has more than 20 years of engineering and project/program management experience and is a licensed professional engineer in the state of Maryland. He has a BS in Mechanical Engineering and MS in Engineering Administration, both from George Washington University. Mr. Ebrahimi comes to us from the NIH Office of Research Facilities Development and Operations where he successfully led a number of highly complex design and construction activities.

In February 2005, **Todd Danielson** joined NLM as Executive Officer. Mr. Danielson comes to us from the National Cancer Institute and brings with him

over 13 years of administrative management experience, most recently as manager of an Administrative Resource Center. Mr. Danielson received his BA in Social Sciences from the University of Pittsburgh and his MS in Business Administration from the Boston University Overseas Program in Heidelberg, Germany. Mr. Danielson joined the NIH as a Presidential Management Intern. At NIH he has assumed positions of increasing responsibility, developing extensive expertise and demonstrating leadership in managing NIH administrative matters.

In February 2005, **Betsy Humphreys** was appointed NLM Deputy Director. Ms. Humphreys received her BA from Smith College and her MLS from the University of Maryland. She was elected a member of the Institute of Medicine and has received numerous other professional honors and awards. During her 30 years at NLM, Ms. Humphreys has served in a variety of critical positions, most recently as the Associate Director for Library Operations. She has provided outstanding leadership and technical direction that resulted in an extensive array of innovative programs. Her numerous exceptional contributions to the Library's programs have been instrumental to NLM's continued success in the development of new information delivery systems, the enhancement of core library services, improved public access to health information, and the Library's increasing role in health data standards.

In February 2005, **Deborah A. Zarin, MD**, joined the LHNCBC's Office of the Director as Assistant Director for Clinical Research Projects. Dr. Zarin will be leading the ClinicalTrials.gov project. Dr. Zarin received her medical degree from Harvard Medical School. She completed her internship in Pediatrics at the Massachusetts General Hospital in Boston and her residency in Psychiatry at the McLean Hospital in Belmont, MA. Prior to joining the LHC, Dr. Zarin was the Director of the Technology Assessment Program at the Agency for Healthcare Quality and Research, DHHS. Dr. Zarin served as Deputy Medical Director and held other important positions at the American Psychiatric Association. She also has held various research, clinical, and teaching positions in the fields of psychiatry and psychiatric epidemiology.

In March 2005, **Yulia Borodina, Ph.D.**, joined the staff of the Computational Biology Branch of the NCBI as a Visiting Fellow (VP). Dr. Borodina received her Master's in medical cybernetics from the Russian State Medical University and her Ph.D. in biochemistry from the Russian Academy of Medical Sciences, Moscow where her research thesis discussed chemical similarity of endogenous and synthetic bioregulators. At

NCBI, Dr. Borodina will undertake algorithm research relevant to PubChem.

In March 2005, **Fabien Fontaine, Ph.D.**, joined the staff of the Computational Biology Branch of the NCBI as a Visiting Fellow (VP). Dr. Fontaine received his Ph.D. in health and life sciences from the Pompeu Fabra University, Barcelona, Spain where he became familiar with cheminformatic methodology. At NCBI, Dr. Fontaine will perform research on algorithms for efficient clustering of compounds by chemical similarity and activity profile.

In March 2005, **Ying (Joanna) He, Ph.D.**, joined the LNCBC. Dr. He received her doctoral degree in Electrical and Computer Engineering from the University of Maryland. For two years Dr. He was a postdoctoral fellow and lab manager in the Distributed Computing and Communications Lab at Colorado State University. Dr. He has a background in mathematical statistics and in designing learning algorithms. At LHNCBC Dr. He is working on Bayesian Learning and Graphical Modeling projects.

In April 2005, **Diane Boehr** was appointed to the position of Head, Cataloging Section, TSD, LO. Ms. Boehr has over 20 years of experience in cataloging, most recently as a supervisor in the Cataloging Section. She has made significant contributions to NLM initiatives including the design of the NLM Archive for electronic resources and metadata schema for NLM's Web resources. Prior to coming to NLM she worked for Costabile Associates on a variety of contracts providing bibliographic services to libraries. Ms. Boehr holds a BS in Chemistry from the City College of New York and a MLS from the University of Maryland, College Park, where she still teaches cataloging.

In April 2005, **Yufeng Jane Tseng, Ph.D.**, joined the staff of the Computational Biology Branch of the NCBI as a Postdoctoral IRTA. Dr. Tseng received her Ph.D. in medicinal chemistry from the University of Illinois at Chicago, where her dissertation focused on the 4D-QSAR Paradigm to Virtual Library Screening and Molecular Similarity Analysis. At NCBI, Dr. Tseng will investigate improved methods for clustering of compounds by bioassay activity profile, and the use of 3D conformers in virtual screening for compounds that may bind to protein targets of known 3D structure for the PubChem database.

In June 2005, **Jessica Fong, Ph.D.**, joined NCBI as a PostDoc from the NIH Undergraduate Scholarship Program. She will be working with Dr. Steve Bryant on evolutionary classification and analysis of protein structures. She recently received her Ph.D. in

computer science from Princeton University. Her thesis described computational techniques for predicting protein interactions via the coiled-coil interaction motif, using protein sequence data. As an intern at AT&T Research Labs during graduate school, she also designed algorithms for network routing and processing massive data sets.

In July 2005, **Brent Bolin** joined the Division of Specialized Information Services as a Health Law Information Specialist after completing the HHS Emerging Leaders Program. Mr. Bolin has a Doctor of Jurisprudence and a Master of Public Affairs from Indiana University. Mr. Bolin conducted research on federal environmental statutes for the U.S. Environmental Protection Agency. During his research fellowship at Indiana, Mr. Bolin contributed legal research towards the publication of a scholarly treatise in the field of public land law. At the SIS, Mr. Bolin will serve as the Project Officer for the Public Health Law Information Project database and will assist in development of various legal information initiatives.

In July 2005, **Timothy Doerr, Ph.D.**, was converted to Staff Scientist after having been an IRTA fellow since 2002. Dr. Doerr received a Ph.D. in theoretical condensed matter physics from Case Western Reserve University and pursued post-doctoral research there on polymers and liquid crystals. He will continue the research he has conducted as an IRTA fellow, namely molecular dynamics simulation of biomolecular interactions, and statistical analysis of pseudogenes and transcription factor binding sites.

In July 2005, **Damir Herman, Ph.D.**, was converted to an IRTA Postdoctoral Fellow after having been a Visiting Fellow for two years. Dr. Herman has a Ph.D. in theoretical condensed-matter physics from Case Western Reserve University. His graduate work was on researching quantum effects in hundreds of electrons confined to mesoscopic structures popularly dubbed quantum dots. His work at NCBI was to use ideas from mesoscopic physics and apply them to noise analysis in microarrays and mutation affinities of transcription factors in yeast, fruit fly, mouse and human.

In July 2005, **Matthew Mailman, Ph.D.**, converted to Staff Scientist at NCBI after one year as a contractor with Computercraft. Dr. Mailman has a Ph.D. in Pathology from Ohio State University, a M.S. in Bioinformatics from the University of Pennsylvania, and a M.S. in Oncology from the University of Wisconsin. His postdoctoral research at the University of Pennsylvania Center for Bioinformatics involved the computational prediction of sequence elements involved in tissue-specific alternative splicing of mRNA. Dr. Mailman will continue working on the development of a database to represent phenotype data and on the development of dbSNP.

In July 2005, **Arthur Petrosian, Ph.D.**, was appointed as Scientific Review Administrator for Extramural Programs. Dr. Petrosian trained in Mathematics, Bioinformatics, and Biomedical Engineering at Moscow State University. He received additional training in biomedical imaging and signal processing at University of Michigan. He was on the faculty of Texas Tech University, gaining the rank of Associate Professor of Neurology. In 2003 he joined NIH's Center for Scientific Resources. While at the Center Dr. Petrosian developed a number of software programs designed to streamline various review processes. At NLM his responsibilities will include administering NLM peer review activities.

In July 2005, **Tao Tao, Ph.D.**, was converted to Staff Scientist. Dr. Tao has been a Kevric contractor since 2000. He earned a Ph.D. in Biomedical Science (Microbiology) from the Medical College of Ohio for work on bacterial restriction-modification system, and did several years of R&D work on PIV genetics and vaccines. He will continue to provide technical support for NCBI resources such as BLAST and Entrez databases as well as custom consultation to researchers within NIH campus.

In August 2005, **Yiming Bao, Ph.D.**, was converted to Staff Scientist after having been a Computercraft contractor since 2001. Dr. Bao earned his Ph.D. in molecular virology from the John Innes Center in Norwich, UK. After three years of postdoctoral training, Yiming became a senior research associate at the Noble Foundation, Ardmore, Oklahoma, where he did research in plant viruses. He will continue working on the NCBI Viral Genome Project.

NLM Associate Fellowship Program

The NLM Associate Fellowship Program is a one-year training fellowship for recent graduates of Masters Degree programs in library and information science. Fellows receive a comprehensive orientation to NLM programs and services during a structured 5-month curriculum phase, and conduct individual projects over the remaining 7-month period. Projects relate to key NLM programs areas and are typically of a research, development, or evaluation nature. Six new Associate Fellows began their year at NLM on September 6, 2005.

Amy Harper received her MLIS in May 2005 from the University of Washington. Her most recent

library experience was with the Program for Appropriate Technology in Health, where she provided reference and document delivery services to staff involved in global health issues. She also has varied experience in academic libraries, including circulation, serials, binding, and acquisitions. Her undergraduate degree is in Latin and Classics.

Mellanye Lackey received her MSI in April 2005 from the University of Michigan. She has experience as a reference assistant in both the medical library and public health library at the University of Michigan. She has also served as a summer intern in Hungary at the Regional Environmental Centre for Central and Eastern Europe. She spent two years in Americorps, serving in the San Francisco Conservation Corps and the Boston public schools. Her undergraduate degree is in Women's Studies.

Lisa Massengale received her MLIS in June 2003 from the University of Pittsburgh. She has two years' professional experience as a reference librarian at Nova Southeastern University in Florida, where the library serves jointly as an academic and public library. In addition to providing reference services, she was involved in library instruction, collection development, and Web page maintenance. She also has prior experience in circulation at a community college. Her undergraduate degree is in English Literature.

Elizabeth Whipple received her MLS in August 2005 from Indiana University. She has experience as a reference assistant in Indiana University's Main Library and Health Sciences Library, where her duties included direct reference, bibliographic instruction, compiling serials data, and Web site content management. She also has experience working with an electronic journals subscriptions system in the Library Information Technology department at Indiana. Her undergraduate degree is in Mathematics.

Oscensio Tom received his MLIS in May 2005 from the University of Arizona. He had a graduate internship in Collection Services at the Arizona Health Sciences Library, where he worked on projects related to gift collections and interlibrary loan. He also has experience as a library assistant in the University's Law Library. As an undergraduate student, he worked with the Community Health Representative Organization of the Navajo Nation, where he developed a data entry procedures manual for a patient registry database and trained Community Health Representatives in its use. His undergraduate degree is in Communication.

Cristina Horta is from Mozambique and is participating in the program as an International Fellow.

She comes to us with 25 years experience as a librarian and manager in Mozambique and Swaziland. Ms. Horta received her MLS degree in 1991 from SUNY–Buffalo, where she also had a graduate assistantship as a reference librarian. She currently is a documentalist for the Ministry of Health's Directorate for Planning and Cooperation. She has been Chief Librarian for the Mananga Agricultural Management Centre in Swaziland, Assistant Librarian (Cataloging) of the University of Swaziland Library, Librarian and Information Officer for The British Council in Mozambique, and Assistant Librarian for the Portuguese Cultural Centre in Mozambique. Her undergraduate degree is in History.

Retirements and Separations

In October 2004, **Jon Retzlaff** departed NLM to become the Director of Legislative Relations at the Federation of American Societies for Experimental Biology. Mr. Retzlaff joined the NLM as Executive Officer in August 2002. As Executive Officer, he provided advice to the Director and other senior staff on administrative management matters and directed the administrative programs including budget, acquisitions, human resources, space management, travel as well as other administrative services.

In December 2004, **Kenneth Niles**, Head, Collection Access Section, retired from NLM after 35 years of Federal government service. Mr. Niles started at NLM in 1991 and he was promoted to Head of Onsite Services, and then in 2001, to Head of the Section. Ken was well known for his enthusiasm and "get the job done" attitude. He made significant improvements to services in the Reading Room, shortening the wait for materials to be paged to under 30 minutes. He often brought in his tools from home to reconfigure workstations or build "Rube-Goldberg" type contraptions that would assist staff in processing requests. In just this past year, Mr. Niles led an effort to improve the interlibrary lending fill rate from 74 percent to 80 percent, and to increase the percentage of requests processed within 12 hours from 80 percent to 96 percent. NLM patrons, staff and colleagues wish Ken all the best in his retirement years.

In January 2005, **Alexa T. McCray, Ph.D.**, retired as Director of the Lister Hill National Center for Biomedical Communications after 19 years of service with NLM. Prior to becoming Director of LHNCBC, Dr. McCray served as the Chief, Cognitive Science Branch. Before joining NLM in 1986, Dr. McCray was a Research Staff Member at IBM's T. J. Watson Research Center. She received her Ph.D. from Georgetown University in 1981, and for three years was

on the faculty there. During her tenure at the NLM, Dr. McCray also served in various organizations, including: member of the Institute of Medicine, a Fellow of the American Association for the Advancement of Science, a Fellow of the American College of Medical Informatics, member of the board of the American Medical Informatics Association, and a past member of the board of the International Medical Informatics Association. Dr. McCray received numerous awards, and, in 2004, ClinicalTrials.gov, a program directed by Dr. McCray, was recognized with the 2004 Innovations in American Government Award by Harvard University.

In May 2005, **Jane Bortnick Griffith**, Assistant Director for Policy Development retired from the Library after 32 years of Federal government service. Prior to joining NLM in 2000, Ms. Bortnick Griffith worked at the Congressional Research Service of the Library of Congress for over 25 years as a Specialist in Information Technology Policy and in senior management positions, providing analytical support to the U.S. Congress on a wide range of issues. She received her BA degree from the University of Wisconsin in 1969 and her MA from Rutgers University in 1970. In her position as Assistant Director for Policy Development, Ms. Bortnick Griffith advised the Director on a variety of policy issues, including access to government information, scientific publishing, intellectual property rights, and privacy. When asked, she also filled essential leadership positions, most recently acting as Deputy Director from August 2004 until January 2005, shepherding the NLM through transitions with formidable analytic skill, outstanding judgment, and an encyclopedic knowledge of science and information policy.

Awards

The 2005 NLM Board of Regents Award for Scholarship or Technical Achievement was awarded to **Frederick B. Wood, DBA** for leadership in the development and implementation of a multidimensional Web evaluation research strategy that has contributed to more effective operation of major NLM Web sites, and spin-off benefits on a trans-NIH basis.

The Frank B. Rogers Award recognizes employees who have made significant contributions to the Library's fundamental operational programs and services. The 2005 awardees were **Lillian R. Kozuma** in recognition of exceptional contributions to the transformation of many of NLM's historical printed bibliographic tools into electronic form, thereby significantly increasing their value through expanded access, and **Frances E. Spina** for outstanding

management of the data creation, maintenance, and quality assurance of citations for MEDLINE, which included phasing out the keyboarding of citations, improving the currency of the data, and increasing the cost effectiveness of the operation.

The NLM Director's Award, presented in recognition of exceptional contributions to the NLM mission, was awarded to three employees: **Jane Bortnick Griffith** for outstanding contributions to policy development and management at the NLM; **Christine C. Ireland** for technical expertise and outstanding management of all NLM Committee Management Activities; and **Ronald F. Stewart** for valuable insight, knowledge, expertise, and leadership spanning 17 years of dedicated service at NLM.

The NIH Merit Award was presented to four employees: **Annmarie A. Carr** for resourcefulness, flexibility, and ingenuity in maintaining administrative efficiency during an extended period of significant changes in personnel practices and services at NIH; **Maria E. Collins** for exceptional management of DOCLINE, NLM's automated interlibrary loan system, and for excellent customer service to over 3,200 medical libraries, world-wide that use this system; **Arnita M. Miles** for carrying out her grant management activities with great efficiency, intelligence, and good cheer during a period of personnel shortage; and **Nadgy P. Roey** for superior leadership and expert guidance managing the NLM's Ethics program and ensuring that the integrity of the program and the public trust were maintained.

The NIH Director's Award was presented to one individual and one group. The individual award recipient was **James R. Marcetich** for leading the expansion of MEDLINE indexing to encompass gene annotation while also increasing indexing productivity, and the group award recipients were **David Lipman, M.D., Ph.D.** and **James Ostell, Ph.D.** for their outstanding contributions to the NIH Public Access Policy Development and Implementation Team.

The Office of the Director, NIH Merit Award was presented to a group that included an NLM employee, **David Landsman, Ph.D.** The citation on the award was in recognition of their work in establishing the NIH Graduate Partnership Program.

The Philip C. Coleman Award recognizes significant contributions to the NLM by individuals who demonstrate outstanding ability to motivate colleagues. The recipient of the 2005 award was **James T. Dean** for his leadership in providing opportunities for improving oneself through educational programs.

The NLM EEO Special Achievement Award was presented to **Dar-Ning Kung, Ph.D.** for his strong leadership, promoting outreach and diversity and enhancing the quality of work life for Asian Pacific American employees.

The U.S. Medicine Frank Brown Berry Prize was presented to **Donald A.B. Lindberg, M.D.** in recognition of outstanding contributions to medicine emanating from the Federal Sector.

The American Medical Women's Association Lila A. Wallis Women's Health Award was presented to **Donald A.B. Lindberg, M.D.** in recognition of exceptional achievements in the successful application of computer technology to health care with notable contributions to information and computer activities in medical diagnosis, artificial intelligence, and educational programs.

Table 14.

FY 2005 Full-Time Equivalents (Actual)

Office of the Director	9
Office of Health Information Programs Development	6
Office of Communications and Public Liaison	8
Office of Administration	39
Office of Computer and Communications Systems	50
Extramural Programs	10
Lister Hill National Center for Biomedical Communications	69
National Center for Biotechnology	140
Specialized Information Services	34
Library Operations	275
TOTAL FTEs	**640**

Board of Regents

The NLM Board of Regents met three times in FY 2005 on February 15–16, May 10–11, and September 20–21. Dr. Thomas Detre, Distinguished Service Professor of Health Sciences, at the University of Pittsburgh, was elected Chairman at the May meeting. Coinciding with the regular Board meetings were meetings of the primary subcommittees including the Extramural Programs Subcommittee, the Subcommittee on Outreach and Public Information, and the Planning Subcommittee. The Chair of each subcommittee provides a summary report to the Board for discussion at each meeting. During 2005, the Board of Regents established two new working groups.

- The Strategic Vision Working Group held its first meeting on April 11–12, 2005 to initiate discussions on a new 2006–2016 NLM long-range plan. Several future long-range planning panels are scheduled to develop this plan and proceed with assessing NLM's mission and current situation, and its potential future contributions to the health and well being of America.

- The Public Access Working Group held its first meeting on July 11, 2005. It was created to advise the Board on the implementation of a new NIH Public Access policy. This policy responds to strong congressional interest in improving the public's access to the published results of NIH-funded research. The policy encourages NIH-supported investigators to submit manuscripts electronically to the NLM's PubMed Central. The working group will review statistical evidence on the impact of the policy, provide suggestions for improving the implementation of the manuscript submission system, assess the extent to which the policy is achieving its stated goals, and suggest any changes to the policy that might further these goals. This working group will meet at least twice a year until 2007, and will prepare a report for presentation to the Board.

Update reports on current programs and projects are provided to the Board members at each meeting by the many NLM divisions. At the February meeting, the Board voted to approve in concept a project to preserve interactive digital publications at the NLM and to expand the Library's research into this area. Each meeting also includes a report by an NLM-funded principal investigator, to provide a sample to the Board of some areas of research funded through the NLM's extramural program. For more detailed information on extramural and grant-related activities of the Board of Regents, see the Extramural Programs section of this report.

NLM Diversity Council

The NLM Diversity Council welcomed four new members in FY 2005: Carmen Aguirre, Maureen Burton, Sue Levine, and Elizabeth Mullen. Each will serve a two-year term. Continuing on the Council are: Patricia Carson, Kathleen Cravedi, Tameka Gore, Donald Jenkins, Melanie Modlin, Helen Ochej, and Bryant Pegram. The Council continues to receive support from its ex-officio members: Todd Danielson, Executive Officer, Mehryar Ebrahimi, Office of Administrative and Management Analysis Services, Pamela Oliver from the NIH Office of Equal Opportunity and Diversity Management, and Nadgy Roey the Program Advisor and Ethics Coordinator for NLM, as well as its distinguished alumni. Patricia

Carson and Melanie Modlin accepted the responsibilities of Council Co-Chairs, Tameka Gore became Council Vice-Chair, and Helen Ochej became Council Secretary.

2005 Accomplishments

NLM Director's Employee Education Fund: The NLM Diversity Council continued its coordination of the NLM Director's Employee Education Fund. In FY 2005, the Fund enabled 50 staff to take 77 classes. The school with the largest number of NLM enrollees was the University of Maryland (29 attendees) with Montgomery College coming in second (8 attendees). Course disciplines enrolled in included: psychology, business, marketing, computer networking, chemistry, economics, mathematics, etc. In addition to traditional classroom instruction, some students took courses on the Internet. The Diversity Council continues its effort to publicize the availability of the fund.

Facility Accessibility and Reasonable Accommodations: The Council continued efforts to upgrade NLM facilities for people with disabilities. Accessibility features in many of the bathrooms in NLM have now been added to accommodate the disabled community. Conference Room B is now equipped with assistive listening devices so that all can participate fully in Conference Room events. The Diversity Council has purchased an electric wheelchair, available upon request for visitors and staff members.

Communication of NLM Diversity: The Diversity Council again collaborated with the Office of Communications and Public Liaison to promote various activities on the NLM Staff Bulletin Board, located outside the cafeteria. This display has proved an excellent setting for celebrating the diversity found at the NLM. The Council voted to have OCPL staffer Fran Sandridge attend meetings on an ex-officio basis, to assist in the design of needed displays.

English Language Courses: The Council continues to support an English-language program, to enable NLM employees to improve their linguistic proficiency speaking and writing English. Following the model used by local literacy programs, the NLM program offers one-on-one tutoring with NLM staff members who volunteer to serve as tutors. Two instructors have received formal training and the course is in place.

School Supply Drive: In 2005, the Diversity Council forged a partnership with Rolling Terrace Elementary School in Takoma Park, Md. The school serves a multi-ethnic population of about 700, including a high percentage of low-income students. In August, NLM staff filled and refilled the collection bins, so that each needy child could have a new backpack loaded with notebooks, pencils, scissors and other back-to-school essentials.

EEO Event: The Library created the first functioning Diversity Council on campus. As part of its effort to ensure that NLM employees know their workplace rights, the Council sponsors various Equal Employment Opportunity events, such as this year's successful "Meet and Greet" session, held June 29th. That session introduced Pam Oliver, EEO Specialist in the NIH Office of Equal Opportunity and Diversity Management. The well-attended talk helped explain NIH's new consolidated EEO structure and familiarized staff with services available in the new EEO Office.

Umbrella Covers: In a move to improve safety and enhance hospitality, the Council has placed disposable plastic umbrella covers at the Library's three principal entrances. The covers prevent rainwater from splashing off umbrellas and onto the floor. They have been well received and well used by employees and Library guests.

Diversity Council Honors NLMers with Awards and Ice Cream Social: The Council sponsored a "Laborless" break from work (right after Labor Day), to honor NLMers whose volunteer activities helped to promote diversity and improve employment opportunities at the Library in 2004–05. Ben and Jerry's supplied the ice cream on the patio adjacent to the Lister Hill Auditorium following the awards ceremony in the auditorium. About 400 NLMers attended. This was the third annual Diversity Council awards ceremony and ice cream social.

Getting to Know NLM, Part 2: After the tremendous success of the series by the same name, which the Diversity Council of 2000–01 created in order to spotlight the work of NLM's many divisions, this year's Council thought it was time for a revival. Under the retooled format, tours of different Library areas like the stacks and the Conservation Lab will be offered, along with talks by employees on such endeavors as NLM-sponsored powwows, international programs and Internet2.

APPENDIX 1:   REGIONAL MEDICAL LIBRARIES

1.   **MIDDLE ATLANTIC REGION**
     The New York Academy of Medicine
     1216 Fifth Avenue
     New York, NY 10029–5283
     (212) 822-7396 FAX (212) 534-7042
     States served: DE, NJ, NY, PA
     URL: http://www.nnlm.nih.gov/mar

2.   **SOUTHEASTERN/ATLANTIC REGION**
     University of Maryland at Baltimore
     Health Science and Human Services Library
     601 Lombard Street
     Baltimore, MD 21201-1583
     (410) 706-2855 FAX (410) 706-0099
     States served: AL, FL, GA, MD, MS, NC,
     SC, TN, VA, WV, DC, VI, PR
     URL: http://www.nnlm.nih.gov/sar

3.   **GREATER MIDWEST REGION**
     University of Illinois at Chicago
     Library of the Health Sciences (M/C 763)
     1750 West Polk Street
     Chicago, IL 60612-7223
     (312) 996-2464 FAX (312) 996-2226
     States served: IA, IL, IN, KY, MI, MN,
     ND, OH, SD, WI
     URL: http://www.nnlm.nih.gov/gmr

4.   **MIDCONTINENTAL REGION**
     University of Utah
     Spencer S. Eccles Health Sciences Library
     10 North 1900 East
     Salt Lake City, Utah 84112-5890
     Phone: (801) 581-8771
     Fax: (801) 581-3632
     States Served: CO, KS, MO, NE, UT, WY
     URL: http://nnlm.gov/mcr

5.   **SOUTH CENTRAL REGION**
     Houston Academy of Medicine Texas Medical
     Center Library
     1133 M.D. Anderson Boulevard
     Houston, TX 77030-2809
     (713) 799-7880 FAX (713) 790-7030
     States served: AR, LA, NM, OK, TX
     URL: http://www.nnlm.nih.gov/scr

6.   **PACIFIC NORTHWEST REGION**
     University of Washington
     Regional Medical Library, HSLIC
     Box 357155
     Seattle, WA 98195-7155
     (206) 543-8262 FAX (206) 543-2469
     States served: AK, ID, MT, OR, WA
     URL: http://www.nnlm.nih.gov/pnr

7.   **PACIFIC SOUTHWEST REGION**
     University of California, Los Angeles
     Louise M. Darling Biomedical Library
     Box 951798
     Los Angeles, CA 90025-1798
     (310) 825-1200 FAX (310) 825-5389
     States served: AZ, CA, HI, NV and U.S.
     Territories in the Pacific Basin
     URL: http://www.nnlm.nih.gov/psr

8.   **NEW ENGLAND REGION**
     University of Massachusetts Medical School
     The Lamar Soutter Library
     55 Lake Avenue, North
     Worcester, MA 01655
     (508) 856-2399 FAX: (508) 856-5039
     States Served: CT, MA, ME, NH, RI, VT
     URL:http://nnlm.gov/ner

APPENDIX 2:   BOARD OF REGENTS

The NLM Board of Regents meets three times a year to consider Library issues and make recommendations to the Secretary of Health and Human Services affecting the Library

Appointed Members:

DETRE, Thomas, M.D. (Chair)
Distinguished Service Prof. of Health Sciences
University of Pittsburgh
Pittsburgh, PA

BUCHANAN, Holly S., Ed. D.
Associate Vice President of Knowledge Management and IT
Health Sciences Library & Informatics Center
University of New Mexico
Albuquerque, NM

CARTER, Ernest L., M.D.
Director, Telehealth Sciences & Advanced Technology Center
Howard University
Washington, D.C.

CHABRAN, Richard, M.L.S., Chair
California Community Technology Policy Group
3081 Sunrise Court
Chino Hills, CA

CONERLY SR., A. Wallace, M.D.
Vice Chancellor Emeritus for Health Affairs
Dean Emeritus, School of Medicine
University of Mississippi
Jackson, MS

GINGRICH, Newt, Ph.D.
Chief Executive Officer
The Gingrich Group
Washington, DC

KARLIS, Vasiliki, D.M.D., M.D.
Associate Professor
Department of Oral and Maxillofacial Surgery
New York University College of Dentistry
New York, NY

Ex Officio Members:

Librarian of Congress

Surgeon General
Public Health Service

Surgeon General
Department of the Air Force

Surgeon General
Department of the Navy

Surgeon General
Department of the Army

Under Secretary for Health
Department of Veterans Affairs

Assistant Director for Biological Sciences
National Science Foundation

Director
National Agricultural Library

Dean
Uniformed Services University of the Health Sciences

APPENDIX 3:   BOARD OF SCIENTIFIC COUNSELORS/ LISTER HILL CENTER

The Board of Scientific Counselors meets periodically to review and make recommendations on the Library's intramural research and development programs.

Members:

FULLER, Sherrilynne S., Ph.D. (Chair)
Professor of Biomedical & Health Informatics
University of Washington School of Medicine
Seattle, WA

CALIFF, Robert M., M.D., Director
Duke Clinical Research Institute
Duke University Medical School
Durham, NC

CARTER, Jerome H., M.D.
Director, Department of Medical Education
Morehouse School of Medicine
Atlanta, GA

CHEN, Hsinchun, Ph.D.
Professor of Management Information Systems
University of Arizona
Tucson, AZ

FERRIN, Thomas E., Ph.D.
Professor of Pharmaceutical Chemistry
University of California
San Francisco, CA

FRIEDMAN, Carol, Ph.D.
Adjunct Professor, Dept. of Medical Informatics
Columbia University
New York, NY

GIUSE, Nunzia B., M.D.
Associate Professor of Biomedical Informatics
Vanderbilt University
Nashville, TN

SRIHARI, Sargur N., Ph.D.
Distinguished Professor
Computer Science & Engineering
State University of NY
Buffalo, NY

APPENDIX 4: BOARD OF SCIENTIFIC COUNSELORS/ NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION

The NCBI Board of Scientific Counselors meets periodically to review and make recommendations on the NLM's biotechnology-related programs.

Members:

PREUSS, Daphne K. Ph.D. (Chair)
Assistant Professor
Molecular Genetics and Cell Biology
University of Chicago
Chicago, IL

FIRE, Andrew Z., Ph.D.
Staff Scientist
Departments of Pathology and Genetics
Stanford University School of Medicine
Stanford, CA

GINSBURG, David, M.D., Professor
Internal Medicine and Human Genetics
Life Sciences Institute
University of Michigan
Ann Arbor, MI

MACKAY, Trudy F., Ph.D.
Professor, Dept. of Genetics
North Carolina State University
Raleigh, NC

SALEMME, F. Raymond, Ph.D.
President
Linguagen Corporation
Cranbury, NJ

SALZBERG, Steven L., Ph.D.
Senior Director of Bioinformatics
University of Maryland
College Park, MD

THOMAS, Annette C., Ph.D.
Managing Director & President
Nature Publishing Group
Macmillan Publishers Ltd.
London, United Kingdom

APPENDIX 5: BIOMEDICAL LIBRARY AND INFORMATICS REVIEW COMMITTEE

The Biomedical Library and Informatics Review Committee meets three times a year to review applications for grants under the Medical Library Assistance Act.

Members:

JENKINS, Carol G., MLS, Director (Chair)
Director, Health Sciences Library
University of North Carolina
Chapel Hill, NC

ARONSKY, Dominik, M.D., Ph.D.
Assistant Professor
Department of Biomedical Informatics
Eskind Biomedical Library
Vanderbilt University Medical Center
Nashville, TN

CAMPBELL, James R., M.D., Professor
Department of Internal Medicine
University of Nebraska Medical Center
Omaha, NE

DUNKER, A. Keith, Ph.D.
Professor of Biochemistry & Molecular Biology
Indiana University School of Informatics
Indianapolis, IN

HUNTER, Lawrence E., Ph.D.
Associate Professor of Pharmacology
University of Colorado Health Sciences Center
Denver, CO

LEHMANN, Harold P., M.D., Ph.D.
Associate Professor, Health Sciences Informatics
Johns Hopkins University
Baltimore, MD

LIDDY, Elizabeth D., Ph.D.
Trustee Professor
School of Information Studies
Syracuse University
Syracuse, NY

NADKARNI, Prakash M., M.D.
Associate Professor, Dept. of Anesthesiology
Center for Medical Informatics
Yale University School of Medicine
New Haven, CT

OGUNYEMI, Omolola I., Ph.D.
Assistant Professor
Department of Radiology
Brigham and Women's Hospital
Boston, MA

PRATT, Wanda, Ph.D.
Assistant Professor
Department of Biomedical & Health Informatics
University of Washington School of Medicine
Seattle, WA

SHEDLOCK, James, AMLS, AHIP, Director
Galter Health Sciences Library
Feinberg School of Medicine
Northwestern University
Chicago, IL

SILVERSTEIN, Jonathan C., M.D.
Assistant Professor of Surgery
University of Chicago
Chicago, IL

SPACKMAN, Kent A., M.D., Ph.D.
Professor of Pathology
Oregon Health and Science University
Portland, OR

TAIRA, Ricky K., Ph.D.
Associate Professor, Dept. of Radiology
University of California
Los Angeles, CA

TANJI, Virginia M., MSLS
Librarian, Health Sciences
John A. Burns School of Medicine
University of Hawaii
Honolulu, HI

TEMPLETON, Etheldra, M.L.S.
Executive Director Library & Info. Systems
Philadelphia College of Osteopathic Medicine
Philadelphia, PA

WALKER, James M., M.D.
Chief Medical Information Officer
Geisinger Health System
Danville, PA

WONG, Stephen T.C., Ph.D.
Associate Professor
Department of Radiology
Brigham and Women's Hospital
Boston, MA

YOKOTE, Gail A., M.S.
Associate University Librarian
Peter J. Shield Library
University of California
Davis, CA

ZHOU, Z. Hong, Ph.D.
Associate Professor of Pathology
University of Texas
Health Science Center – Medical School
Houston, TX

APPENDIX 6: LITERATURE SELECTION TECHNICAL REVIEW COMMITTEE

The Literature Selection Technical Review Committee meets three times a year to select journals for indexing in MEDLINE.

Members:

MCCLURE, Lucretia W., M.A. (Chair)
Special Assistant to the Director
Countway Library of Medicine
Harvard Medical School
Boston, MA

BAUCHNER, Howard, M.D., Professor
Professor of Pediatrics
Boston Medical Center
Boston, MA

DELCLOS, George L., M.D.
Professor of Environmental & Occupational Health
University of Texas Health Science Center
Houston, TX

FREY, John J., M.D.
Professor and Chair
Department of Family Medicine
University of Wisconsin
Madison, WI

KAPLAN, Jerry, Ph.D.
Professor of Pathology
University of Utah School of Medicine
Salt Lake City, UT

LI, Rong, Ph.D.
Associate Professor
Stowers Institute for Medical Research
Kansas City, MO

MANNING, Phil, M.D.
Professor of Medicine Emeritus
(University of Southern California)
Corona del Mar, CA

RACZ, Gabor B., M.D.
Grover Murray Professor & Director of Pain Services
Department of Anesthesiology
Texas Tech University Health Sciences Center
Lubbock, TX

SHARPS, Phyllis W., Ph.D.
Associate Professor
School of Nursing
Johns Hopkins University
Baltimore, MD

SIEGEL, Vivian, Ph.D.
Department of Pathology
University of California
San Francisco, CA

SOEHNER, Catherine B., M.L.S.
Head, Science & Engineering Library
University of California
Santa Cruz, CA

SPANN, Melvin, Ph.D.
Chairman & CEO
SGH Medtech Consult, Inc.
Silver Spring, MD

STERNBERG, Esther M., M.D.
Director, Integrative Neural Immune Program
National Institute of Mental Health
Bethesda, MD

TOM-ORME, Lillian, Ph.D.
Research Assistant Professor
Dept. of Family and Preventive Medicine
University of Utah
Salt Lake City, UT

VAN PEENEN, Hubert J., M.D.
Professor of Pathology (Retired)
Ponca City, OK

APPENDIX 7: PUBMED CENTRAL NATIONAL ADVISORY COMMITTEE

The PubMed Central National Advisory Committee meets twice a year to review and make recommendations about PubMed Central.

WILLIAMS, James F., M.S. (chair)
Dean of Libraries
University of Colorado
Boulder, CO

BAKER, Shirley K., M.A.
Dean and Vice Chancellor
Libraries and Information Technology
Washington University
St. Louis, MO

DELAMOTHE, Anthony P., M.D.
Editor, British Medical Journal
London, England

EISEN, Michael B
Genome Sciences
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA

JOSEPH, Heather D., M.A.
President & Chief Operating Officer
BioOne
Washington, DC

KAPLAN, Samuel, Ph.D.
Professor and Chair
Microbiology and Molecular Genetics
University of Texas Health Science Ctr.
Houston Medical School
Houston, TX

KAUFMAN, Paula T., M.B.A.
University Librarian
University of Illinois at Urbana–Champaign
Urbana, IL

KILEY, Robert J., M.S.
Head, Systems Strategy, Wellcome Library

Wellcome Trust
London, England

KIRSCHNER, Marc W., Ph.D.
Professor and Chair
Department of Systems Biology
Harvard Medical School
Boston, MA

LAPPIN, Debra R., J.D.
President
Princeton Partners Ltd.
Englewood, CO

ROEHR, Robert J., M.S.
Journalist
Washington, D.C.

RUBIN, Gerald M., Ph.D.
Investigator
Dept. of Molecular & Cellular Biology
University of California
Berkeley, CA

RYAN, Mary L., M.P.H., M.L.S
Director, UAMS Library
University of Arkansas for Medical Sciences
Little Rock, AR

SO, Anthony D., M.D., Director
Program on Global Health and Technology Access
Terry Sanford Institute of Public Policy
Duke University
Durham, NC

VARKI, Ajit P., M.D.
Professor of Cellular and Molecular Medicine
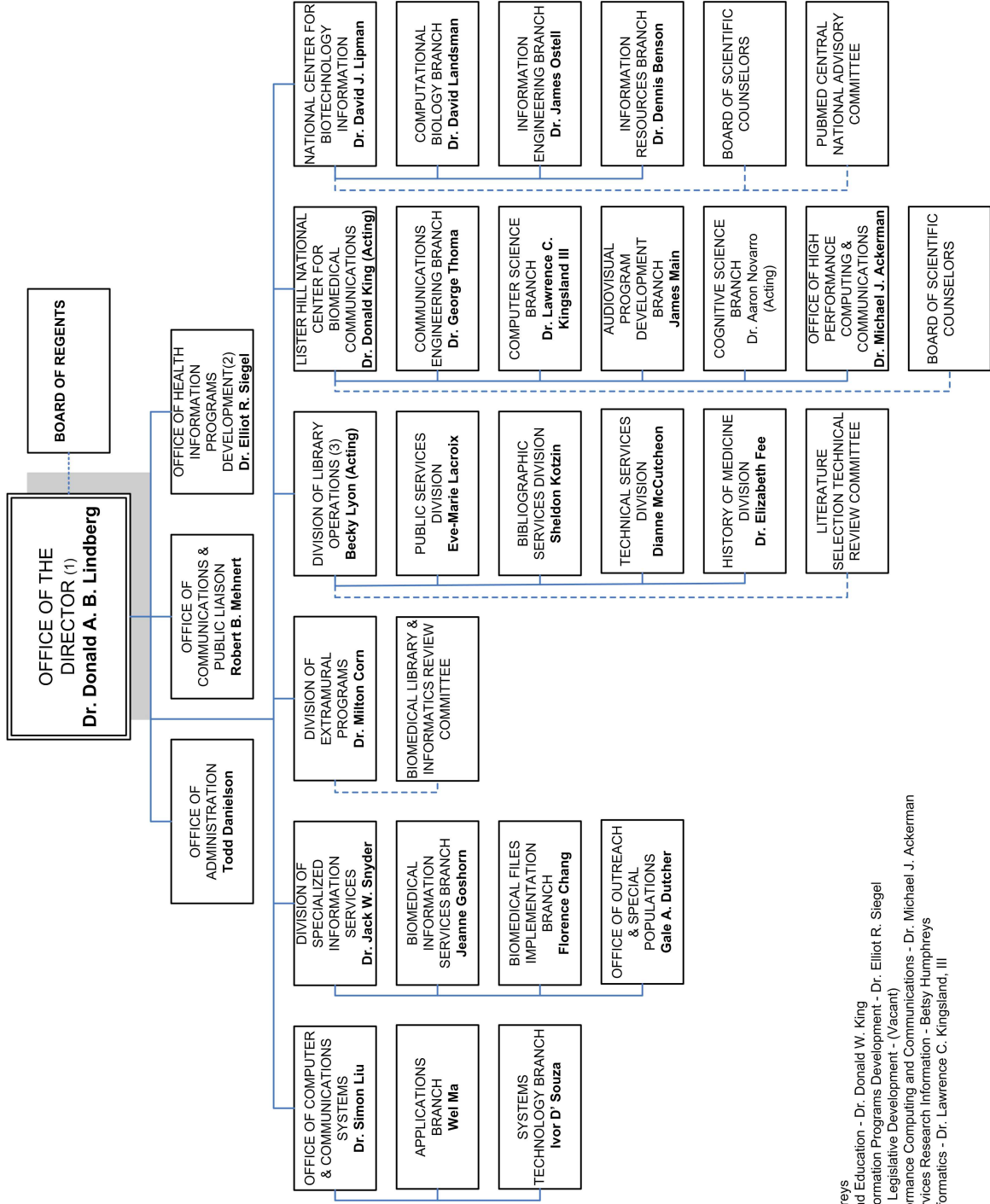University of California, San Diego
La Jolla, CA

APPENDIX 8: ORGANIZATIONAL ACRONYMS AND INITIALISMS USED IN THIS REPORT

AAHSL	Association of Academic Health Sciences Libraries
ACP	American College of Physicians
ACSI	American Customer Satisfaction Index
AHRQ	Agency for Healthcare Research and Quality
ALTBIB	Alternatives to Animal Testing
AME	Automated Metadata Extraction
AMIA	American Medical Informatics Association
AMPA	American Medical Publishers Association
AMWA	American Medical Women's Association
APDB	Audiovisual Program Development Branch
BISTI	Biomedical Information Science and Technology Initiative
BLAST	Basic Local Alignment Search Tool
BLIRC	Biomedical Library and Informatics Review Committee
BOR	Board of Regents
BSD	Bibliographic Services Division
BSN	Bioinformatics Support Network
CAS	Collection Access Section
CBIR	Content-Based Image Retrieval
CCB	Configuration Control Board
CCDS	Consensus CoDing Sequence
CCRIS	Chemical Carcinogenesis Research Information System
CDD	Conserved Domain Database
CEB	Communications Engineering Branch
CgSB	Cognitive Science Branch
ChemIDplus	Chemical Identification File
CIT	Center for Information Technology
CoreBio	Core Bioinformatics Facility
CPT	Current Procedural Terminology
CRISP	Computer Retrieval of Information on Scientific Projects
CSB	Computer Science Branch
CSI	Commission on Systemic Interoperability
CSR	Center for Scientific Review
CT	Computer Tomography
DART	Developmental and Reproductive Toxicology
dbMHC	Database for the Major Histocompatability Complex
DDBJ	DNA Data Bank of Japan
DCMS	Data Creation and Maintenance System
DEAS	Division of Extramural Administrative Support
DHHS	Department of Health and Human Services

DIRLINE	Directory of Information Resources Online
DTD	Document Type Definition
EBI	European Bioinformatics Institute
EEO	Equal Employment Opportunity
EFTS	Electronic Funds Transfer Service
EMBL	European Molecular Biology Laboratory
EMS	Emergency Medical Services
EP	Extramural Programs
EPA	Environmental Protection Agency
eRA	Electronic Research Administration
EST	Expressed Sequence Tag
ETIC	Environmental Teratology Information Center
FDA	Food and Drug Administration
FIC	Fogarty International Center
FNLM	Friends of the National Library of Medicine
FTE	Full Time Employee
Gbps	Gigabits per Second
GDS	GEO DataSet
GEO	Gene Expression Omnibus
GENSAT	Gene Expression Nervous System Atlas
GHR	Genetics Home Reference
GIS	Geographic Information System
GPS	Global Positioning System
GSA	General Services Administration
GSS	Genome Survey Sequences
GUI	Graphic User Interface
HapMap	Haplotype Map
HAVnet	Haptic Audio Video Network for Education Technology
HBCU	Historically Black Colleges and Universities
HHS	Health and Human Services
HIPAA	Health Insurance Portability and Accounting Act
HLA	Human Leukocyte Antigen
HL7	Health Level Seven, Inc.
HMD	History of Medicine Division
HSDB	Hazardous Substances Data Bank
HPCC	High Performance Computing and Communications
HRSA	Health Resources and Services Administration
HSRProj	Health Services Research Projects
HSRInfo	Health Services Research Information
HSRR	Health Services and Sciences Research Resources

HSTAT	Health Services and Technology Assessment Text	NHGRI	National Human Genome Research Institute
IAIMS	Integrated Advanced Information Management Systems	NHLBI	National Heart, Lung, and Blood Institute
ICMJE	International Committee of Medical Journal Editors	NIA	National Institute on Aging
ICs	Institutes and Centers (of NIH)	NIAID	National Institute of Allergy and Infectious Diseases
IHM	Images from the History of Medicine	NIBIB	National Institute of Biomedical Imaging and Bioengineering
ILL	Interlibrary Loan	NICHSR	National Information Center on Health Services Research and Health Care Technology
ILS	Integrated Library System		
IRIS	Integrated Risk Information System	NIDCD	National Institute on Deafness and other Communication Disorders
IT	Information Technology		
ITER	International Toxicity Estimates for Risk	NIDDK	National Institute of Diabetes, Digestive, and Kidney Diseases
ITK	Insight Toolkit		
JDI	Journal Descriptor Indexing	NIEHS	National Institute of Environmental Health Sciences
LACT	Drugs and Lactation (database)		
LAN	Local Area Network	NIGMS	National Institute of General Medical Sciences
LHC	Lister Hill Center		
LHNCBC	Lister Hill National Center for Biomedical Communications	NIH	National Institutes of Health
		NINDS	National Institute of Neurological Disorders and Stroke
LO	Library Operations		
LOINC	Logical Observations: Identifiers, Names, Codes	NIOSH	National Institute for Occupational Safety and Health
		NIST	National Institute of Standards and Technology
LSTRC	Literature Selection Technical Review Committee		
		NLM	National Library of Medicine
MARG	Medical Article Records Groundtruth	NLP	Natural Language Processing system
MARS	Medical Article Records System	NN/LM	National Network of Libraries of Medicine
MDoT	MEDLINE Database on Tap	NNO	National Network Office
MDT	Multimedia Database Tool	NOSC	Network Operations and Security Center
MEDLARS	Medical Literature Analysis and Retrieval System	NOVA	National Online Volumetric Archive
		NRCBL	National Reference Center for Bioethics Literature
MEDLINE	MEDLARS Online		
MeSH	Medical Subject Headings	NSF	National Science Foundation
MHC	Major Histocompatability Complex	NTCC	National Online Training Center and Clearinghouse
MIM	Multilateral Initiative on Malaria		
MIRS	Medical Information Retrieval System	OAM	Office of Administrative Management
MLA	Medical Library Association	OCCS	Office of Computer and Communications Systems
MLAA	Medical Library Assistance Act		
MMDB	Molecular Modeling DataBase	OCHD	Coordinating Committee on Outreach, Consumer Health and Health Disparities
MMS	MEDLARS Management Section		
MMTx	MetaMap Technology Transfer		
MTI	Medical Text Indexer	OCPL	Office of Communications and Public Liaison
MTMS	MeSH Translation Management System		
NCBC	National Centers for Biomedical Computing	OCR	Optical Character Recognition
		OD	Office of the Director
NCBI	National Center for Biotechnology Information	OHIPD	Office of Health Information Programs Development
		OMB	Office of Management and Budget
NCCS	NIH Consolidated Collocation Site	OMIA	Online Inheritance in Animals (database)
NCI	National Cancer Institute		
NCRR	National Center for Research Resources	OMIM	Online Mendelian Inheritance in Man (database)
NCVHS	National Committee on Vital and Health Statistics		
		OMSSA	Open Mass Spectrometry Search Algorithm
NEI	National Eye Institute	OSIRIS	Open Source Independent Review and Interpretation System
NGI	Next Generation Internet		
NHANES	National Heath and Nutrition Examination Surveys	PAHO	Pan American Health Organization

PCA	Personal Computer Advisory Committee	SOAP	Simple Object Oriented Protocol
PCR	Polymerase Chain Reaction	SPER	System for the Preservation of Electronic Resources
PDA	Personal Digital Assistant		
PDR	Publisher Data Review	SII	Scalable Information Infrastructure
PDB	Protein Data Bank	STB	Systems Technology Branch
PDF	Portable Document Format	STTR	Small Business Technology Transfer Research
PDR	Publisher Data Review		
PHLI	Public Health Law Information Project	STS	Sequence Tagged Site
PHS	Public Health Service	TEHIP	Toxicology and Environmental Health Information Program
PLAWARe	Programmable Layered Architecture With Artistic Rendering		
		TIE	Telemedicine Information Exchange
PMC	PubMedCentral	TILE	Text to Image Linking Engine
PRS	Protocol Registration System	TIOP	Toxicology Information Outreach Project
PSD	Public Services Division	TOXLINE	Toxicology Information Online
RCSB	Research Collaboratory for Structural Bioinformatics	TOXNET	Toxicology Data Network
		TPA	Third Party Annotation (database)
RefSeq	Reference Sequence (database)	TRI	The Toxics Release Inventory
RFA	Request for Applications	TSD	Technical Services Division
RFP	Request for Proposals	TTP	Turning the Pages
RML	Regional Medical Library	UMLS	Unified Medical Language System
RNAi	RNA interference	UPS	Uninterrupted Power Supply
RRF	Rich Release Format	VAST	Vector Alignment Search Tool
RTECS	Registry of Toxic Effects of Chemical Substances	VHP	Visible Human Project
		WebMIRS	Web-based Medical Information Retrieval System
SBIR	Small Business Innovation Research		
SEF	Serials Extract File	Web-STOC	Web-Services Technology Operations Center
SIDA	Swedish International Development Agency		
		WGS	Whole Genome Shotgun
SII	Scalable Information Infrastructure	WIISARD	Wireless Internet Information System for Medical Response in Disasters
SIS	Specialized Information Services		
SMART	Scalable Medical Alert and Response Technology	WISER	Wireless Information System for Emergency Responders
SNOMEDCT	Systematized Nomenclature of Medicine Clinical Terms	XML	Extensible Markup Language

National Library of Medicine

BOARD OF REGENTS

OFFICE OF THE DIRECTOR (1)
Dr. Donald A. B. Lindberg

OFFICE OF ADMINISTRATION
Todd Danielson

OFFICE OF COMMUNICATIONS & PUBLIC LIAISON
Robert B. Mehnert

OFFICE OF HEALTH INFORMATION PROGRAMS DEVELOPMENT (2)
Dr. Elliot R. Siegel

OFFICE OF COMPUTER & COMMUNICATIONS SYSTEMS
Dr. Simon Liu

- APPLICATIONS BRANCH **Wel Ma**
- SYSTEMS TECHNOLOGY BRANCH **Ivor D' Souza**

DIVISION OF SPECIALIZED INFORMATION SERVICES
Dr. Jack W. Snyder

- BIOMEDICAL INFORMATION SERVICES BRANCH **Jeanne Goshorn**
- BIOMEDICAL FILES IMPLEMENTATION BRANCH **Florence Chang**
- OFFICE OF OUTREACH & SPECIAL POPULATIONS **Gale A. Dutcher**

DIVISION OF EXTRAMURAL PROGRAMS
Dr. Milton Corn

- BIOMEDICAL LIBRARY & INFORMATICS REVIEW COMMITTEE

DIVISION OF LIBRARY OPERATIONS (3)
Becky Lyon (Acting)

- PUBLIC SERVICES DIVISION **Eve-Marie Lacroix**
- BIBLIOGRAPHIC SERVICES DIVISION **Sheldon Kotzin**
- TECHNICAL SERVICES DIVISION **Dianne McCutcheon**
- HISTORY OF MEDICINE DIVISION **Dr. Elizabeth Fee**
- LITERATURE SELECTION TECHNICAL REVIEW COMMITTEE

LISTER HILL NATIONAL CENTER FOR BIOMEDICAL COMMUNICATIONS
Dr. Donald King (Acting)

- COMMUNICATIONS ENGINEERING BRANCH **Dr. George Thoma**
- COMPUTER SCIENCE BRANCH **Dr. Lawrence C. Kingsland III**
- AUDIOVISUAL PROGRAM DEVELOPMENT BRANCH **James Main**
- COGNITIVE SCIENCE BRANCH Dr. Aaron Novarro (Acting)
- OFFICE OF HIGH PERFORMANCE COMPUTING & COMMUNICATIONS **Dr. Michael J. Ackerman**
- BOARD OF SCIENTIFIC COUNSELORS

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION
Dr. David J. Lipman

- COMPUTATIONAL BIOLOGY BRANCH **Dr. David Landsman**
- INFORMATION ENGINEERING BRANCH **Dr. James Ostell**
- INFORMATION RESOURCES BRANCH **Dr. Dennis Benson**
- BOARD OF SCIENTIFIC COUNSELORS
- PUBMED CENTRAL NATIONAL ADVISORY COMMITTEE

---

1. Deputy Director – Betsy Humphreys
   Deputy Director for Research and Education - Dr. Donald W. King
   Associate Director for Health Information Programs Development - Dr. Elliot R. Siegel
   Assistant Director for Policy and Legislative Development - (Vacant)
   Assistant Director for High Performance Computing and Communications - Dr. Michael J. Ackerman
   Assistant Director for Health Services Research Information - Betsy Humphreys
   Assistant Director for Applied Informatics - Dr. Lawrence C. Kingsland, III

2. Includes International Programs

3. Includes:

   National Network of Libraries of Medicine, Head - Dr. Angela Ruffin
   Medical Subject Headings Section, Chief - Dr. Stuart Nelson
   National Information Center on Health Services Research & Health Care Technology, Head - Marjorie A. Cahn