

**DEPARTMENT OF HEALTH AND HUMAN SERVICES
NATIONAL INSTITUTES OF HEALTH
NATIONAL LIBRARY OF MEDICINE**

**MINUTES OF THE BOARD OF REGENTS
February 15-16, 2005**

The 138th meeting of the Board of Regents was convened on February 15-16, 2005, at 9:00 a.m. in the NLM Board Room, Building 38, National Library of Medicine (NLM), National Institutes of Health (NIH), Bethesda, Maryland. The meeting was open to the public from 9:00 a.m. to 4:30 p.m., followed by a closed session for consideration of grant applications until 5:00 p.m. On February 16, the meeting was reopened to the public from 9:00 a.m. until adjournment at 11:45 a.m.

MEMBERS PRESENT:

Dr. Holly Buchanan, University of New Mexico
Dr. Ernest Carter, Howard University
Dr. A. Wallace Conerly, University of Mississippi Medical Center
Dr. Richard Dean, Wake Forest University
Dr. Thomas Detre, University of Pittsburgh
The Honorable Newt Gingrich, The Gingrich Group
Dr. William Stead [Chair], Vanderbilt University

EX OFFICIO AND ALTERNATE MEMBERS PRESENT:

Ms. Eleanor Frierson, U.S. Department of Agriculture
MGEN Joseph E. Kelley, U.S. Department of the Air Force
Dr. Deanna Marcum, U.S. Library of Congress
Dr. Michael Pazzani, National Science Foundation
Col. John Powers, U.S. Department of the Army
Dr. Vernon Schinski, Uniformed Services University of the Health Sciences
Ms. Mary Ann Tatman, U.S. Department of Veterans Affairs
Admiral Carol Turner, U.S. Department of the Navy

CONSULTANTS TO THE BOR PRESENT:

Dr. H. Kenneth Walker, Emory University School of Medicine

SPEAKERS AND INVITED GUESTS PRESENT:

Dr. Richard Carmona, Surgeon General, Department of Health and Human Services
Dr. George Hripcsak, Columbia University
Ms. Helena M. VonVille, University of Texas School of Public Health at Houston
Dr. Elias Zerhouni, Director, National Institutes of Health

MEMBERS OF THE PUBLIC PRESENT:

Ms. Jill Helm, Allscripts
Mrs. Mary Lindberg
Ms. Linda Lingle, Friends of the NLM
Ms. Anne Michaels, Friends of the NLM

February 15-16, 2005 - Board of Regents

Ms. Amy Pearman, The Gingrich Group

FEDERAL EMPLOYEES PRESENT:

Dr. Donald A.B. Lindberg, Director, NLM

Ms. Betsy Humphreys, Deputy Director, NLM

Dr. Michael Ackerman, High Performance Computing & Communications, NLM

Ms. Joyce Backus, Division of Library Operations, NLM

Mr. Jeff Beck, National Center for Biotechnology Information, NLM

Mr. Dennis Benson, National Center for Biotechnology Information, NLM

Ms. Susan Buyer, Office of Health Information Program Development, NLM

Mr. Charlie Cook, National Center for Biotechnology Information, NLM

Dr. Milton Corn, Division of Extramural Programs, NLM

Mr. Todd Danielson, Executive Office, NLM

Mr. Ivor D'Souza, Office of Computer and Communications Systems, NLM

Ms. Gale Dutcher, Division of Specialized Information Services, NLM

Mr. Mehryar Ebrahimi, Office of Administrative and Management Analysis Services, NLM

Dr. Valerie Florance, Division of Extramural Programs, NLM

Dr. Charles Friedman, Division of Extramural Programs, NLM

Ms. Jane Bortnick Griffith, Office of the Director, NLM

Ms. Wendy Hadfield, Executive Office, NLM

Mr. Tim Hayes, Office of Extramural Research, NIH

Ms. Dana Haza, Office of the Director, NLM

Mr. Walter Hickel, Division of Library Operations, NLM

Ms. Colette Hochstein, Division of Specialized Information Services, NLM

Ms. Christine Ireland, Division of Extramural Programs, NLM

Dr. Donald King, Office of the Director, NLM

Dr. Dar-Ning Kung, Lister Hill Center, NLM

Mr. Chris Krahe, Division of Specialized Information Services, NLM

Ms. Michelle Krever, Division of Extramural Programs, NLM

Ms. Eve-Marie Lacroix, Division of Library Operations, NLM

Ms. Janet Laylor, Office of the Director, NLM

Mr. Daniel Lerman, Office of the Director, NIH

Dr. David Lipman, National Center for Biotechnology Information, NLM

Dr. Simon Liu, Office of Computer and Communications Systems, NLM

Mr. Howard Lu, Lister Hill Center, NLM

Dr. Robert Logan, Lister Hill Center, NLM

Ms. Becky Lyon, Division of Library Operations, NLM

Ms. Christie Moffatt, Division of Library Operations, NLM

Ms. Dianne McCutcheon, Division of Library Operations, NLM

Mr. Robert Mehnert, Office of Communication and Public Liaison, NLM

Mr. Dwight Mowery, Division of Extramural Programs, NLM

Mr. Michael Muin, Lister Hill Center, NLM

Dr. Aaron Navarro, Lister Hill Center, NLM

February 15-16, 2005 - Board of Regents

Dr. Stuart Nelson, Medical Subject Headings Section, NLM
Ms. Julia Royall, Office of Health Information Program Development, NLM
Dr. Angela Ruffin, Division of Library Operations, NLM
Mr. Scott Sidel, Office of Computer and Communication Systems, NLM
Dr. Elliot Siegel, Office of Health Information Program Development, NLM
Dr. Hua-Chuan Sim, Division of Extramural Programs, NLM
Dr. Jack Snyder, Division of Specialized Information Services, NLM
Ms. Marti Szczur, Division of Specialized Information Services, NLM
Dr. Paul Theerman, Division of Library Operations, NLM
Dr. George Thoma, Lister Hill Center, NLM
Dr. Deborah Zarin, Lister Hill Center, NLM

I. OPENING REMARKS

Dr. William Stead, Chair of the NLM Board of Regents, welcomed the Regents, alternates, consultants, and guests to the 138th meeting of the Board. He announced that on December 17, 2004, the President of Georgia declared that NLM consultant and former Regent Dr. Kenneth Walker was an honorary citizen of that nation. He was cited for “his contributions to the creation and implementation of the Georgia–American Medical Partnership Project.”

II. REPORT FROM THE SURGEON GENERAL, PHS

Surgeon General Richard Carmona thanked the National Library of Medicine for its role as a strong partner with the Office of the Surgeon General in its efforts to advance and protect the health and well-being of the American people. He said that he has three priorities: prevention, public health preparedness, and eliminating health disparities. Woven throughout this agenda is the issue of health literacy. Dr. Carmona described several projects related to improving America’s health literacy, including the launch of the Surgeon General’s Report on Bone Health and Osteoporosis, and the Surgeon General’s Family History Initiative, a multi-agency project. He said that 2005 is “The Year of the Healthy Child.” It will embrace the most comprehensive agenda ever set forth by a Surgeon General for a single year. It will include all aspects of child health, starting with prenatal care and going through all the developmental stages of childhood and adolescence. On February 21 he will participate in “Birth Day Live,” a program on the Discovery Health Channel that will feature live deliveries as an opportunity to communicate with the American public about child health. Dr. Carmona said he hopes the NLM will partner with his office in making The Year of the Healthy Child a powerful force for improving the life of our children. In response to a question from Dr. Stead about how to gain acceptance for the idea of the Electronic Health Record (EHR), Dr. Carmona said that this subject also is connected to the need for improved health literacy. He said that the President is pushing to move the EHR forward; now we have to package the concept in a culturally competent manner and demonstrate to the public that in fact the EHR is of vital importance to healthcare.

III. FUTURE MEETING DATES

The Board of Regents will meet next on May 10–11, 2005. The Board meeting next fall is on September 20-21. The dates of February 7–8, 2006, were adopted for the following meeting.

IV. SEPTEMBER 2004 MINUTES OF MEETING

The Regents approved without change the minutes from the September 21–22, 2004 meeting.

V. NLM DIRECTOR'S REPORT

Dr. Donald Lindberg reported that NLM closed out FY 2004, expending \$318,365,000 of the \$318,560,000 appropriation. The appropriation was supplemented with \$11,101,000 from other government agencies. NLM's FY 2005 budget is \$323,346,000. This is 2.1% over the FY 2004 level. Dr. Lindberg drew the Board's attention to Congressional language in the House and Senate budget reports that encourage the Library's efforts in outreach, PubMed Central, minority health, access to research results, and the need for expanded facilities. In the area of personnel, Dr. Lindberg said that Betsy Humphreys, former Associate Director for Library Operations, has been appointed NLM Deputy Director. He thanked Jane Griffith for taking on that role in an acting capacity since last summer. He also announced that Mr. Todd Danielson has been appointed NLM Executive Officer, replacing Mr. Jon Retzlaff, who resigned from Federal Service. Other new appointments are Mr. Mehryar Ebrahimi as Chief of the Office of Administrative Management and Dianne McCutcheon as Chief of the Technical Services Division. Last December the Library lost Dr. Alexa McCray who retired as Director of the Lister Hill National Center for Biomedical Communications and Mr. Ken Niles who retired as Head of the Collection Access Section. Dr. Donald King, Acting Director of the Lister Hill Center, introduced to the Board several new staff members: Dr. Deborah Zarin, who will lead the ClinicalTrials.gov project and Mr. Michael Muin, who will be working on an online tool that facilitates the practice of evidence-based medicine. The Director reported that the Board will hear tomorrow from the Subcommittee on Planning about developing a new Long Range Plan for the NLM.

On last December 2–3, the Library and the National Network of Libraries of Medicine co-sponsored a Symposium on Community-Based Health Information Outreach. The participants compared notes about which health information strategies work and which don't work. Former Board Chair Eugenie Prime delivered the keynote address. Dr. Lindberg said that clinical trials (and the database ClinicalTrials.gov that NLM created and maintains for NIH) are the object of interest by the Congress. There are many who believe that we need a system that registers *all* clinical trials. The editors of several major medical journals simultaneously published an editorial saying they will require the registration of clinical trials prior to consideration of manuscripts for publication. In addition to trials sponsored by government agencies and drug companies, NLM has also agreed to add international trials to ClinicalTrials.gov, citing a "national competent authority" in each country from which trials are registered. Dr. Lindberg

February 15-16, 2005 - Board of Regents

said that at the upcoming March 15 American Medical Publishers' Association, AMPA and the NLM are sponsoring a symposium, "Changing Models and Changing Relationships in Publishing." A public event at the University of Missouri in December launched the MedlinePlus "Go Local" in that state. Missouri joins North Carolina in making local health and medical services within the state easily accessible to users of MedlinePlus. More states will be joining the system this spring. Arrangements are being made to have the major NLM exhibition, "Changing the Face of Medicine," travel to some 60 sites around the country. As an adjunct to that exhibition, in cooperation with the American Medical Women's Association, NLM has introduced a "Local Legends" gallery on the Web that features more than 200 women physicians around the U.S. nominated by their Congressional representatives. Dr. Lindberg reported that through the intercession of Regent Newt Gingrich, Dr. Julie Gerberding, Director of the Centers for Disease Control and Prevention invited Dr. Lindberg and several NLM officials to discuss how the two organizations can work closely together. A number of good ideas for collaboration resulted from the meeting. Yesterday, Dr. Lindberg said, the papers of Francis Crick were added to NLM's "Profiles in Science" Web site. The Regents will hear about the site later in the meeting.

Dr. Lindberg ended his presentation by discussing the concept of the "interactive publication" of the future. Interactive scientific publications will be able to combine the traditional features of print and online publication with the full capabilities of computer-based information systems. The result will be "papers" with which readers can interact in order to see further than the current limited range of illustrations and static summaries of data. He showed several typical articles and discussed how the reader might reasonably wish to see much more, or different, data and images than those provided. NLM's Profiles in Science, with its combination of text, images, and video has some of the elements of an interactive publication. Dr. Lindberg also showed features from several medical journals he found that have interactive-publication features. We need to study interactive publications to find out if they facilitate understanding and learning. Since NLM is committed to acquire and preserve such publications we need to look into how best to do this. We hope to be able to find publishers to work with on the subject of interactive publication. Following Dr. Lindberg's presentation, there was a general discussion by the Regents about the potential of interactive publication. Speaker Gingrich wondered "what would a fully digitized American health experience be?" As long as the markets (practitioners, teachers, students, public) are kept in silos, the capital investment to get the most out of our knowledge will never be made. We are "right at the edge of being able to change everything," he said. Dr. Lindberg said that the presentation scheduled for tomorrow by Dr. Lipman will give a taste of this interconnectedness. There was discussion about the reluctance of medical schools to cooperate in matters of curriculum and to achieve a critical mass that would result in change. Speaker Gingrich said that if NLM were to reach out and suggest a conference co-hosted by NLM, Google, Microsoft, and others, and invite all the deans of the medical schools—that might be a very interesting one-day interaction.

VI. TOXMAP – ENVIRONMENTAL HEALTH E-MAPS

Ms. Marti Szczur, Associate Director for Specialized Information Services (SIS), introduced the subject of TOXMAP, a new Geographic Information System created by the SIS. The Division of Specialized Information Services operates a Toxicology and Environmental Health Information Program that traces its establishment at the NLM back to the 1960s. Several of the products and databases of that program, such as the Household Products Database, Tox Town, and the Wireless Information System for Emergency Responders (WISER) have been demonstrated to the Regents over the last few years. Ms. Szczur said that the new TOXMAP system (<http://toxmap.nlm.nih.gov>) relies on three TOXNET components: the Hazardous Substances Data Bank, TOXLINE (a bibliographic file), and the Toxic Release Inventory (TRI—information from the EPA). TRI is the only one of the three with geo-coded data. She briefly described how TOXMAP was developed and how focus groups were involved in improving it before release. Ms. Szczur then went online and demonstrated various features of the system, which is intended for a broad audience, including biomedical researchers, public health departments, health sciences libraries, public health faculties, community health programs, and the general public. Feedback so far from users has been very positive.

Following Ms. Szczur's presentation, Ms. Helena VonVille, Library Director, University of Texas School of Public Health in Houston, described how she has been using TOXMAP since last June in her courses and workshops on the utilization of localized health data in Texas. Before that time she didn't have a way to get to good local information on toxic releases. She was introduced to TOXMAP at a medical library conference last May and, beginning in June, was able to demonstrate it in courses and workshops. The first workshop involving the public, "The Environment & You: An Environmental Hazards Workshop," was held last week and was featured in an article in the *Houston Chronicle*. Ms. VonVille said it would be useful to include sociodemographic information in TOXMAP, connecting release data to population characteristics. Also, it would be helpful to be able to see releases by year, rather than just cumulative data, and to rank the substance displays with those having more serious consequences at the top. She made several additional suggestions for useful changes in TOXMAP. Ms. VonVille said that a big advantage of TOXMAP is that it was designed for both the researcher and the public looking for information about environmental pollutants in their community. It is "an incredibly important service" that provides a valuable research tool with unbiased information for faculty, public health students, and the general population.

Following Ms. VonVille's presentation, Ms. Szczur presented candidate next steps in the evolution of TOXMAP: linking to the EPA's National Priority List, adding transportation data (airports, etc.), linking to the Coast Guard National Response Center incident reports, adding occupational code information from the Census Bureau, providing numerous sources of health data (mortality and disease atlas maps), displaying multiple chemicals on a single map, and being able to localize displays even more. Dr. Carter commented that being able to get this kind of information about a local community is tremendously useful, and to be able to overlay it with information about chronic diseases, for example, would make it a very powerful tool. Dr. Walker

February 15-16, 2005 - Board of Regents

said that the progress made by NLM's Toxicology and Environmental Health Information Program in the last decade is tremendous. The improvement in usability of its various databanks and information resources is notable. Speaker Gingrich suggested that software that provides ZIP code level data might be applicable to TOXMAP. Dr. Dean asked whether other information could be overlaid on the maps, for example, salt intake and fast food consumption. Ms. Szczur said that the infrastructure is there so that this could be done. Speaker Gingrich suggested that NLM might talk to the National Geospatial-Intelligence Agency whose specialty is creating information flow about geographic points. Much of their information is not classified.

VII. FRANCIS CRICK "PROFILES IN SCIENCE" LAUNCH

The National Library of Medicine (in Bethesda) and the Wellcome Library for the History and Understanding in Science (in London) established a live video link in the Board of Regents Room to launch the introduction into Profiles in Science an extensive selection from the papers of Nobelist Frances Crick. Dr. Lindberg said that this is a wonderful collaboration between the two institutions. Francis Crick is a giant in his field and he led the revolution in molecular biology in the 50's and 60's. This transatlantic collaboration makes sense, since Crick worked both in the U.K. and the U.S. Dr. Lindberg thanked the Wellcome staff who made it possible for the NLM Profiles in Science to have access to the Crick papers. From the London end of the video link, Clare Matterson, Director of the Wellcome Library's Medicine, History and Science Division, reciprocated in the name of her team of librarians, archivists, and digital imaging specialists who worked on the Crick project with words of thanks and congratulations to NLM for the successful launch of the Crick Web site. Following these remarks, Dr. Walter Hickel, a historian with the Digital Manuscripts Program, History of Medicine Division, and Christie Moffatt, Digital Manuscripts Program Manager, demonstrated some of the highlights from the Crick papers on Profiles in Science (<http://profiles.nlm.nih.gov/SC/>).

VIII. PRESERVATION OF DIGITAL INFORMATION

Dr. Deanna Marcum, Associate Librarian for Library Services, Library of Congress, reported to the Board about the LC's National Digital Information Infrastructure and Preservation Program (NDIIPP). She said it was gratifying to know that the NLM was willing to take "one difficult chunk of this really difficult problem." In the print world, we know how to preserve information and long-term preservation is a responsibility of the national libraries and big research libraries. Electronic information is different: we don't "own" the materials as we do with print, we license access to it. One great difficulty in preserving digital information is getting the rights from the creators and publishers to allow us to do something more with the information. The National Digital Information Infrastructure and Preservation Program at the Library of Congress was created to ensure access over time to a rich body of digital content through the establishment of a national network of committed partners. The goals of the program are to develop a national collection and preservation strategy; build a network of partners; develop a technical architecture that supports a network; and foster research into new tools, techniques and models for digital preservation. The program was established by legislation in December 2000. Up to \$175 million

February 15-16, 2005 - Board of Regents

is potentially available with matching funds. A National Digital Strategy Advisory Board was established composed of librarians, technologists, academics, legal experts, creators/artists, and publishers. Dr. Lindberg is a member and Betsy Humphreys and others at NLM have also been involved. A report from the Board was submitted to and approved by Congress in 2002: “Preserving Our Digital Heritage.” Now the hard work of implementing the plan begins. There are three areas of focus: a network of preservation partners, technical preservation architecture, and basic digital preservation research. Dr. Marcus described each in some detail. The 100 institutions that are part of the partnership program met in January; affinity groups have been created to discuss cross-cutting concerns and shared concerns and their reports will be made in three years.

Following Dr. Marcum’s presentation, Ms. Becky Lyon, NLM Acting Associate Director for Library Operations, said that the Library’s Long Range Plan calls for NLM to take a leadership role in developing a program to ensure permanent access to electronic information in biomedicine, and to work with other national organizations to advance the digital preservation agenda more broadly. The Library is currently active in journal preservation by way of the PubMed Central archive (the Board will hear more about this tomorrow). Ms. Lyon briefly described the “DTD”—the Document Type Definition developed by the NCBI—that in an XML document (either journal or book) provides a list of the elements, attributes, comments, notes, and entities contained in the document. Because the preservation of digital resources offers a rich field for active research in computer science and engineering, NLM’s Lister Hill Center has been much involved. They have concentrated on two key functions: automated metadata extraction (the only way large collections can be handled) and file migration (useful for converting files in formats that face obsolescence). Ms. Lyon then introduced Margaret Byrnes, Head of NLM’s Preservation and Collection Management Section, who described how NLM determined which of its Web documents would be kept permanently available. She defined (and gave examples of) the four permanence ratings: Permanence Not Guaranteed; Permanent—Dynamic Content; Permanent—Stable Content; and Permanent—Unchanging Content. Ms. Byrnes said that NLM has developed an online archive on its main Web site for outdated or superseded permanent Web documents. Current work involves archiving non-HTML files, archiving resources from all areas of NLM, and working with other libraries to identify their permanence rating needs. Mr. Lyon next described how NLM is collaborating with other organizations: federal agencies, publishers, universities, libraries, and private companies. She said they have determined that additional R&D is required—both within NLM and by other organizations—to support the preservation of interactive digital publications. She ended by saying that the NLM requests Board approval to expand the Library’s research into the new area of interactive publications—both their development and preservation.

Eleanor Frierson commented that these presentations show how far we have come in this important area. NLM’s permanence ratings and the NCBI toolkit are both welcome and will help us deal with stable material. How to deal with volatile material is more problematic, she said. The TOXMAP demonstration this morning is an example of such volatile material. She believes that the recommendation before the Board is highly desirable. She also offered to have the

February 15-16, 2005 - Board of Regents

National Agricultural Library collaborate with NLM on the research work that is being proposed. Dr. Holly Buchanan said that she supported the suggestion made this morning by Speaker Gingrich that NLM co-host a conference that would look at a fully digitized interactive medical “publication.” With the success of NLM’s Building Symposium and Outreach Symposium as models, perhaps such a conference would be effective in creating these diverse partnerships. Dr. Buchanan moved that the Board approve in concept NLM’s new research objectives and work in digital preservation, as presented by Ms. Lyon. Dr. Detre commented about security—the need to “preserve” materials in more than one place so that material would not be lost in a disaster. Dr. Marcum said that the Library of Congress has an alternate computer facility in another location to forestall such an eventuality. Dr. Lindberg said this was also the case with NLM. Speaker Gingrich asked what happens to a digitized world if the electricity grid gets seriously interrupted? This is a strategic challenge. The Board approved unanimously the motion made by Dr. Buchanan.

IX. NLM AND HEALTH DATA INTEROPERABILITY STANDARDS

Betsy Humphreys, NLM Deputy Director, said that electronic health data standards are a key component of the national health information network, needed for efficient health care, research, public health, and emergency detection and response. Underlying NLM’s interest in health data standards is the assumption that Electronic Health Records (EHR) will make it easier to deliver relevant information at the time and place important decisions are made. “Interoperability standards” is a broad term: NLM’s interest is in the subset that deals with data content, standard vocabularies (for use in messages and records), and mapping between clinical vocabularies and administrative code sets. NLM’s current Long Range Plan commits the Library to work with others to establish, maintain, test, and promote the use of such standards. NLM is also committed to using its Unified Medical Language System (UMLS) resources as a vehicle for uniform distribution of administrative coding systems, detailed clinical vocabulary, and mappings between the two. Ms. Humphreys briefly recounted the 15-year history and many reports on the subject of efforts to establish standards. There was substantial agreement among those reports in what needed to be done to establish, coordinate, improve, and promote use of standards. History will probably show that the 1996 HIPAA (Health Insurance Portability and Accounting Act) had a major positive impact on the development of health data standards. Ms. Humphreys briefly noted the influence of HIPAA and its relationship with and influence on the National Committee on Vital and Health Statistics, the Consolidated Health Informatics project (a cross-agency eGov initiative), and the Medicare Modernization Act that requires e-prescribing standards and established a Commission on Systemic Interoperability.

In 1996, NLM was asked to assist in selecting the code sets and classifications to be used in HIPAA standard transactions. As previously reported to the Board, NLM supports the ongoing development of LOINC (for lab tests/instrument observations). NLM has also developed the RxNorm for clinical vocabulary and has negotiated the U.S.-wide license agreement for SNOMED CT, a broad-based clinical terminology. NLM’s UMLS Metathesaurus serves as the common distribution channel for the HIPAA code sets and other medical vocabularies. One of its

February 15-16, 2005 - Board of Regents

main purposes is to connect all the synonyms for concepts from different vocabularies. The current version of the Metathesaurus contains one million concepts and almost five million concept names from 130 different vocabularies. Recent UMLS enhancements have improved the technical foundation for the next step in the process to achieving U.S. health data standardization, namely, to coordinate the development of selected standards so they become a non-overlapping, interlocking set. Former HHS Secretary Tommy Thompson designated NLM as the central HHS coordinating body for standard clinical terminology. This coordination will involve (1) eliminating duplication of effort and establishing explicit relationships between vocabularies, (2) ensuring that the standard clinical vocabularies line up with medical messaging and health record formats, and (3) mapping between standard clinical vocabularies and administrative code sets and other important vocabularies. Ms. Humphreys briefly discussed several examples of mapping, involving LOINC, RxNorm, and SNOMED CT. NLM is assisting HHS by serving as the operational home of the Commission on Systemic Interoperability, created to develop a comprehensive strategy for the adoption and use of interoperability standards. NLM is likely to have responsibility for expending some of the funds from the Office of the National Coordinator of Health Information Technology (Dr. David Brailer). NLM is also working to broaden participation in standards development and to promote use and feedback—one example is working other NIH components to promote testing and use of standards in the clinical research enterprise. We look forward to reporting some of the positive effects of standardization at future Board meetings.

Following Ms. Humphreys' presentation, Dr. Stead introduced Ms. Dana Haza, the Director of the Commission on Systemic Interoperability. She has experience in both health care operations and policy formulation in the public and private sectors. Ms. Haza briefly described the makeup of the Commission. Speaker Gingrich commented on the absence of an Electronic Health Record by saying that "every day that we fail to get this done we kill people." This should be done without delay, even knowing that any initial system will have to mutate. In California, Kaiser Permanente and Sutter, two hospital systems, have started using Epic, the same commercial Electronic Health Record system. Unfortunately (as of two months ago) the two systems cannot talk to each other. With just a few additions and changes, virtually all of Northern California could be covered. Speaker Gingrich said that, as a first step, he would accept an electronic version of a paper health record as being superior to the paper record itself. If we modify Stark Anti-kickback regulations "every hospital in the country will be able to provide Electronic Health Record capability to every doctor in the country." As to funding, he said that in a multi-trillion dollar federal budget "they can find 100 million dollars for anything they want." Since all medication transactions are required to be electronic, he suggested that we could develop an "Electronic Drug Record" as the first step toward an Electronic Health Record, and expert systems could soon be checking interactions for every medication. Additional capabilities could be grafted onto this system.

Dr. Stead agreed with Speaker Gingrich's notion of concentrating on providing access to the health record in whatever form it has been captured in parallel with solving the complexity of information systems. The latter is hard because historically we have captured information with

different definitions in the internal systems of a facility. Before it can be interchanged, this information has to be mapped to a standard—a human process. The translations are not precise, and you increase the imprecision as the information is passed on. Putting standards between systems is needed, but it doesn't get us to where we want to be, which is to have either humans or computers able to accurately represent the meaning of the content. Dr. Stead said that the key is getting controlled terminology plus computable terms that include standard definitions; if the NLM can be the source of such technologies we will have a breakthrough. SNOMED CT is the first giant step in that direction. Since this standardization process takes time, we need a two-prong track: NLM to take responsibility for that role (as it has begun to) and then do what Speaker Gingrich said—the simple things that will make information accessible in whatever form it is now.

X. DISCOVERING AND APPLYING KNOWLEDGE IN CLINICAL DATABASES

Dr. George Hripcsak, Professor and Vice Chair of the Department of Biomedical Informatics, Columbia University, also serves as the Chair of NLM's Biomedical Library and Informatics Review Committee. He described his NLM-funded project, "Discovering and Applying Knowledge in Clinical Databases," as a data-mining project that employs technology that achieves accuracy and efficiency that could, in principle, enable nationwide screening. Imagine, he said, having a tool like a research assistant that could read every message and note in the national health information infrastructure—millions of admit notes, progress notes, discharge summaries, outpatient notes, etc.—and answer any question you have about them. The goal of the tool is to be able to answer complex clinical questions for a large population. Today we can ask simple questions for large populations and complex questions for small populations. What we need is a detailed, clinically relevant view of patients in electronic form. Dr. Hripcsak described the clinical data repository Columbia has been collecting for 15 years on 2.5 million patients. Much of it is coded data—claims data, laboratory data, etc. Many of these data elements would be essential items in an Electronic Health Record. However the claims data is frequently inaccurate, incomplete, or misleading and the laboratory and medication data need context. The narrative data in the repository is useful, but searching it with precision is difficult. Dr. Hripcsak described the work they have done in natural language processing (a system called MedLEE) which can translate narrative data into codes (via the UMLS). He said the results of tests showed that the automated system achieved "expert-like" performance, doing just as well as the doctor. He presented a number of scenarios of how MedLEE could read and parse the text. Once this database exists, Dr. Hripcsak said, we can use data mining techniques to derive useful knowledge from the data. He gave several examples of "machine learning" that let the machine work with a clinical algorithm to discover new knowledge, for example to determine when a patient should be admitted. The MedLEE and data mining system were also validated against 60,000 discharge summaries in NYPORTS (New York Patient Occurrence Reporting and Tracking System). There was a .9996 specificity rate, that is, of 10,000 cases that didn't have a disease, only four were falsely picked. Dr. Hripcsak said that if every hospital sent in their discharge summaries (over 30 million a year, all in electronic form) they could be added to the MedLEE data system and serve as a tremendous source of information for researchers and

February 15-16, 2005 - Board of Regents

clinicians. He gave several scenarios of how the national database could be usefully queried. Dr. Hripcsak concluded that a clinical data repository is “an amazing resource.” There are 2.5 million patients in their repository, but there is no reason the number couldn’t be 300 million in a nationwide data resource.

Following Dr. Hripcsak’s presentation, Mary Ann Tatman said the work he described is especially significant because with today’s interest in the Electronic Health Record and data repositories, to be able to read the data and translate it into something that improves clinical outcomes is very important. She asked whether data in that repository could be used to formulate alerts. Dr. Hripcsak said that the need for tuberculosis isolation (one of his examples) was formulated as an alert. It has also been used to identify patients who are candidates for clinical trials. In response to another question, he said that MedLEE has been shared on an experimental basis with several other institutions. Dr. Conerly asked whether MedLEE was in routine use. Dr. Hripcsak said that the TB and clinical trial applications were both routine uses of the database. Dr. Lindberg commented that the system focuses on medication errors; what about other types of errors? Dr. Hripcsak said that using NYPORTS and the 45 different adverse events it reports, the goal of MedLEE is to track a variety of errors in diagnosis and failed procedures.

XI. EXTRAMURAL PROGRAMS REPORT

The Board of Regents unanimously approved the Extramural Programs Operating Procedures for 2005. They were unchanged from last year. The FY 2005 EP budget was reviewed briefly with explanation of relationship between recent doubling of budget, current flat budget, and relatively lower success rate anticipated for 2005.

Dr. Corn explained the Responsibilities of the Board for compliance with the 2005 Biennial Report: “Compliance with NIH Policy on Inclusion Guidelines.” This report is required to describe the manner in which the NLM has complied with the statute concerning the appropriate inclusion of women and minorities in extramural activities. Dr. Corn summarized the NLM policy and experience in 2004. Although we are not supporting any formal clinical trials, many of the projects we fund include human subjects. NLM follows all applicable NIH regulations for inclusion of women, children, and minorities. Inclusion considerations are a specific evaluation point during the peer review of each application. Discussion is open to the entire review committee in the presence of the Scientific Review Administrator and other staff of the Extramural Programs Division. Violations of inclusion requirements can result in refusal to review, or tabling of review until further information obtained, or instruction to staff to withhold any award until NLM is satisfied that applicant has met.

The Regents were informed that NLM will initiate Internet Assisted Review (IAR) for the Biomedical Library and Informatics Review Committee meeting scheduled for March 10–11, 2005. The pros and cons of IAR were discussed as well as the three-phase process. NLM intends to permit the “edit phase” for most of its reviews.

February 15-16, 2005 - Board of Regents

Outreach is an important element of NLM's mission and is carried out by a variety of programs from several NLM Divisions. EP Outreach activities were described: Outreach goals are the focus of a significant fraction of the Connection and Translational Informatics grants, but outreach projects are not uncommon in Information Systems grants, some aspects of the Integrated Advanced Information Management Systems program, and from time to time are the subject of biomedical informatics research grants. Outreach activities supported by NLM grant programs were summarized with an overview of grants and expenditures during past five years. Examples of Outreach projects were described.

To address the need for public health informatics specialists, the NLM and the Robert Wood Johnson Foundation (RWJF) have agreed to collaborate in the creation of specialized public health training tracks within some of NLM's 18 currently funded training sites. A limited subset of these training sites will be selected to receive supplemental funding explicitly for training in public health informatics. The preponderance of these supplemental funds will support the trainees themselves, but each selected site will have the option of allocating some resources to faculty support and/or program development. RWJF has provided \$4 million to support this initiative for four years, in the form of a grant to the Foundation for NIH. RWJF and the NLM will be partners in the planning and implementation of the program, including the selection of the programs to receive funding in the form of supplements to their training grants.

MEETING CLOSED FOR THE REVIEW OF GRANT APPLICATIONS February 16, 2005, 4:30 P.M.

XII. COMBATING CYBER ATTACKS

Dr. Simon Liu, Director of NLM's Office of Computer and Communications Systems, reported to the Board about how NLM's extensive computer operation is dealing with threats, exposures, and vulnerabilities to cyber attacks. He used the analogy of dealing with the flu—we are dealing with something with a recurring, invisible, widespread threat, to which certain segments of the population are especially vulnerable. The formula describing this is $\text{Risk} = \text{Threat} * \text{Vulnerability} * \text{Exposure}$. Dr. Liu said the threat consists of more than 68,000 viruses/worms that have been identified, and the 350 new ones that are created every month. Although we can't control the threats, we *can* control our vulnerability and exposure to the threats. He described several exposures that are unavoidable—connections to the outside world. As to vulnerabilities, we apply patches as soon as they are identified; more than 3,000 have been identified to date. An unprotected machine will survive an average of 17 minutes before a cyber attack. Dr. Liu described the community's experience with the recent "SoBig" computer worm and its variants, and he presented some statistics about the number and kind of attacks that have been detected by NLM intrusion detection systems. He also discussed the "race" between the development of a virus or worm by a hacker and the creation of a patch by software vendors to protect against it. The window of time for remediation, unfortunately, is shrinking, as hackers get more sophisticated. Dr. Liu said that "defense in depth" is our best protective strategy. He showed how NLM had a number of mechanisms to control our vulnerability and exposure:

February 15-16, 2005 - Board of Regents

partitioned network, multilevel firewalls, closed port policy, intrusion detection systems, content screening, and host detection and audit. NLM cooperates closely with other security sites operated by government agencies, universities, and vendors. NLM also systematically scans its operations regularly for vulnerabilities and takes steps to fix any found. Dr. Liu concluded by summarizing: attacks will continue to increase; vulnerabilities are rising; reactive windows are narrowing; and a layered proactive approach is the best line of defense.

Dr. Thomas Detre asked about whether our cyber defense could withstand a concerted attack by other countries. Dr. Liu said that is unknown. Our best approach is not response, but engineering our infrastructure so as to be able to anticipate and deal with such events in a layered defense.

Dr. Lindberg commented that most of the agencies and companies that have suffered virus attacks have never put on the available patches. NLM is more careful in this regard.

XIII. REPORT FROM THE NIH DIRECTOR

NIH Director Dr. Elias Zerhouni said that there are synergies that we need to invest in across diseases and across the 27 Institutes and Centers at NIH. The links among diseases are apparent now as they never were before—the human genome contributed much to this. We are engaging in a proactive dialog across NIH trying to identify the gaps and opportunities and, most important, identify what it is no single Institute can do on its own that the NIH *must* do. This is what we call the “Roadmap.” He said that it is obvious that the cost of healthcare is fundamentally driven by our focus over the past 100 years on intervening once disease has struck. What is important today is to find ways of intervening 20 years before the disease (diabetes, heart disease for example) strikes. Having better early knowledge about the development of disease processes, and intervening earlier, is orders of magnitude more cost effective. At the heart of this approach is information. This is why we support many of the initiatives of NLM. Interoperable databases and software systems, not just systems that present static data, are a first step. Our analysis shows that we are going to deal with a tissue of intelligent databases that will have the ability to connect, not at the technical level, but at the informational level. The recently announced PubChem, part of the NIH Roadmap initiative, is an open database for small molecules that you can in an intelligent manner link to any source of medical research information. It is the cumulative nature of the information that gives it power. Dr. Zerhouni gave a few examples of how PubChem might be used. The set of tools in Entrez, created by NLM’s National Center for Biotechnology Information (NCBI), is unique in the world. The recently announced public access policy is a “transforming event” in this process, he said. Never before have scientists had a venue where they can voluntarily put NIH-supported, peer-reviewed, published data. NLM, through PubMed Central, should become an enhancement to the world of information exchange in science. Dr. Zerhouni described to the Board the new policy, which goes into effect May 2, 2005. In addition to serving as a public access vehicle, the new system will be a powerful tool for NIH to manage its portfolio. Dr. Zerhouni praised Dr. Lipman, NCBI Director, as a “visionary leader”; many of the ideas now being discussed originated in discussion with him. Some of the publishers, likewise, are forward looking. *Nature*, for example, has created a publishing venue site for data from NIH grantees in the area of cell

February 15-16, 2005 - Board of Regents

signaling, a fundamental subsystem of biology. Dr. Zerhouni suggested that the Board of Regents consider creating a new Working Group to oversee the public access venture.

Following his presentation, there was a discussion of the role of peer review in the new public access system. In response to a question concerning the need for additional space for the Library, Dr. Zerhouni was not optimistic about the near-term prospects in today's difficult budget milieu. Chairman Gingrich said that we have to think through how we explain in public what a "library" is in an age of digital information. The term "library" is a turn-off to many in the Congress. After the morning's Planning Subcommittee meeting, he said he was convinced that there should be a "National Library of Science" parallel to the National Library of Medicine. The nonmedical sciences do not have an integrative database system comparable to that created by the NLM. Dr. Zerhouni agreed and said that the new PubChem system is a bridge between two of the sciences. A motion was made and seconded to create the Working Group to oversee the public access system suggested by Dr. Zerhouni. The motion was passed. Dr. Thomas Detre has agreed to serve as Chair; Dr. Deanna Marcum will also serve on the Working Group.

XIV. ARCHIVING NIH RESEARCH IN PUBMED CENTRAL

Dr. David Lipman, Director of the National Center for Biotechnology Information, discussed the background of the new public access proposal from NIH. The announcement of the policy coincides with significant changes in scientific publishing, driven by, on the one hand, massive investments in information technology by individuals and institutions and, on the other, new high-throughput methods in biomedical research that greatly increase the amount of data associated with publications. The associated data not only allows the user to examine the data in depth, but to be able to extend the data to make new discoveries. He said that utility, usage, cost, and permanence are four drivers of change when discussing the archiving of research results. Dr. Lipman described how the utility of a publication can be maximized by allowing a diversity of access to the complete data associated with a paper. Similarly, maximizing free and unrestricted usage (by PubMed, for example) and minimizing the delay in availability of the data, increases the value of the information to science, education (including high school and undergraduate education), and the general public. He then briefly discussed digital archiving and its challenges, and the respective responsibilities of publishers, libraries, and the NLM.

Dr. Lipman demonstrated PubMed Central, showing how a scientist can use the database to find information, for example, on the genetics of late onset Alzheimer's disease. NCBI converts the content of articles into a consistently structured set of documents to facilitate programs extracting the content and linking it to factual databases. Dr. Lipman showed how a variety of information connected to late onset Alzheimer's mapping could be found. He then demonstrated in PubChem how detailed chemical information, including structural images, could be located and displayed. Statistics on PubMed Central show that there are 160 journals that are voluntarily participating and there are a total of 350,000 items entered into PubMed Central. This will grow rapidly when all NIH-supported papers begin to be added and as material from back issues of journals are digitized. At the present time about 2 million users access the database monthly. Dr.

February 15-16, 2005 - Board of Regents

Lipman ran through a sample scenario of how an NIH-supported author would submit his or her manuscript to PubMed Central after it is accepted for journal publication, and the steps the NLM would take to add it to the database. The new Working Group, which will report to the BOR, will assist in evaluating the implementation of Public Access, e.g., its timeliness, level of participation by grantees and its use by the research community. Dr. Lipman said that if there were 100 percent participation, PubMed Central would be receiving about 200 papers a working day.

In response to a question from Dr. Michael Pazzani, Dr. Lipman said that NIH's start date is May 2, 2005 and only manuscripts accepted after that date will be included. Receipt of earlier material has not been budgeted for. Chairman Gingrich commented that the public access policy has the potential to redefine what constitutes traditional electronic publication. He could foresee building tools that would communicate the research in lay language. He also suggested that an economic analysis be done to quantify the cost-savings to society by the policy's reduction of research costs and acceleration of new discoveries through more effective use of data. He ventured an opinion that the net result would be 'brehtaking'. Dr. Stead said that this last point, the "magnifier effect of the Library," is a subject that should be fleshed out in the NLM Long Range Plan.

XV. REPORT FROM THE LONG-RANGE PLANNING SUBCOMMITTEE

Dr. William Stead reported on the subcommittee's meeting (February 16, 2005). The focus of discussion was on convening a small meeting as a first step in a planning process that would bring together key people with ideas about what "a fully digital world would look like." What can NLM do that would accelerate the change in that direction? NLM staff will undertake to arrange such a meeting.

XVI. REPORT FROM THE SUBCOMMITTEE ON OUTREACH AND PUBLIC INFORMATION

Dr. Richard Dean briefly reported on the three major topics that were discussed at yesterday's Subcommittee meeting: the Information Rx project, the Go-Local initiative, and the Local Legends program that is a part of the "Changing the Face of Medicine" exhibition and Web site.

XVII. ADJOURNMENT

The meeting was adjourned at 11:45 a.m.

February 15-16, 2005 - Board of Regents

ACTIONS TAKEN BY THE BOARD OF REGENTS:

- Approval of the September 21-22, 2004 Board of Regents Minutes
- Approval of February 7-8, 2006 Meeting Dates
- Concept approval – Preservation of Digital Information (Interactive Publications) Project
- Approval of 2005 Board Operating Procedures
- EP Subcommittee Concurrence and En Bloc Grant Approval
- Establishment of Public Access Working Group

I certify that, to the best of my knowledge, the foregoing minutes and attachments are accurate and complete.

Donald A.B. Lindberg, M.D.
Director, National Library of Medicine

William W. Stead, M.D.
Chair, NLM Board of Regents