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

MINUTES OF THE BOARD OF REGENTS
September 18-19, 2007

The 146th meeting of the Board of Regents was convened on September 18-19, 2007, at 9:00 a.m. in the Board Room, Building 38, National Library of Medicine (NLM), National Institutes of Health (NIH), in Bethesda, Maryland. The meeting was open to the public from 9:00 a.m. to 2:40 p.m., followed by a closed session for consideration of grant applications until 3:00 p.m. On September 19, the meeting was reopened to the public from 9:00 a.m. until adjournment at 12:00 p.m.

MEMBERS PRESENT [Appendix A]**:**
Dr. Cynthia Morton [Chair], Brigham and Women's Hospital
Mr. Richard Chabran, California Community Technology Policy Group
Dr. John Connolly, University of California, Irvine
Dr. Carol Friedman, Columbia University
Dr. C. Martin Harris, The Cleveland Clinic Foundation
Dr. O. Wayne Isom, New York Presbyterian-Weill Cornell Medical School
Dr. Louis Rossiter, The College of William and Mary
Ms. Eileen Stanley, Ecolab, Inc.

EX OFFICIO AND ALTERNATE MEMBERS PRESENT:
Ms. Eleanor Frierson, U.S. Department of Agriculture
Ms. Gail Graham, U.S. Department of Veterans Affairs
Dr. Haym Hirsh, National Science Foundation
Major General Thomas Loftus, U.S. Department of the Air Force
Dr. Patrick Malone, U.S. Department of the Navy
Ms. Kathryn Mendenhall, Library of Congress
Rear Admiral Helena Mishoe, Office of the Surgeon General, PHS
Col. John Powers, U.S. Department of the Army
Dr. Dale Smith, Uniformed Services University of the Health Sciences

CONSULTANTS TO THE BOR PRESENT:
Dr. Marion Ball, Johns Hopkins School of Nursing/IBM Research
Dr. Holly Buchanan, University of New Mexico
Dr. Thomas Detre, University of Pittsburgh
Dr. H. Kenneth Walker, Emory University School of Medicine

SPEAKERS AND INVITED GUESTS PRESENT:
Dr. Margaret Humphreys, Duke University
Mr. Jorge Lambrinos, Roybal Institute

MEMBERS OF THE PUBLIC PRESENT:
Mrs. Mary Lindberg
Dr. William Hole, Retired NLM

FEDERAL EMPLOYEES PRESENT:

Dr. Donald A.B. Lindberg, Director, NLM
Ms. Betsy Humphreys, Deputy Director, NLM
Dr. Donald King, Deputy Director for Research and Education, NLM
Dr. Michael Ackerman, High Performance Computing & Communication, NLM
Ms. Grace Ajuwon, Associate Program, NLM
Ms. Suzanne Aubuchon, Office of the Director, NLM
Ms. Joyce Backus, Division of Library Operations, NLM
Ms. Susan Buyer, Office of Planning & Analysis, NLM
Dr. Olivier Bodenreider, Cognitive Science Branch, NLM
Ms. Lisa Boyd, Office of the Director, NLM
Mr. Allen Browne, Cognitive Science Branch, NLM
Ms. Sarena Burgess, Associate Program, NLM
Mr. Michael Chung, Communication Engineering Branch, NLM
Ms. Kathy Cravedi, Office of Communications & Public Liaison, NLM
Dr. Milton Corn, Division of Extramural Programs, NLM
Ms. Celeste Dade-Vinson, Office of the Director, NLM
Mr. Todd Danielson, Executive Office, NLM
Ms. Robin Dupuis, Budget Reporting Legislative Branch, OD
Ms. Gale Dutcher, Division of Specialized Information Services, NLM
Dr. Elizabeth Fee, History of Medicine Division, NLM
Dr. Marcelo Fiszman, Cognitive Science Branch, NLM
Ms. Kate Flewelling, Associate Program, NLM
Dr. Valerie Florance, Division of Extramural Programs, NLM
Dr. Zoe Huang, Division of Extramural Programs, NLM
Ms. Christine Ireland, Division of Extramural Programs, NLM
Ms. Joanna Karpinski, Associate Program, NLM
Ms. Paula Kitendaugh, Reference and Web Services Section, NLM
Mr. Sheldon Kotzin, Division of Library Operations, NLM
Ms. Brenda Linares, Association Program, NLM
Dr. David Lipman, National Center for Biotechnology Information, NLM
Dr. Robert Logan, Lister Hill Center, NLM
Dr. Teri Manolio, National Human Genome Research Institute, NIH
Dr. Clement McDonald, Lister Hill Center, NLM
Ms. Melanie Modlin, Office of Communications & Public Liaison, NLM
Mr. David Nash, Office of the Director, NLM
Dr. Aaron Navarro, Lister Hill Center, NLM
Mr. Michael North, Division of Library Operations, NLM
Dr. Elaine Ostrander, National Human Genome Research Institute, NIH
Dr. Glen Pearson, Communications Engineering Branch, NLM
Dr. Steven Phillips, Office of the Director, NLM
Ms. Shana Potash, Office of Communications & Public Liaison, NLM

Ms. Melissa Resnick, Associate Program, NLM
Dr. Thomas Rindflesch, Cognitive Science Branch, LHC, NLM
Dr. Graciela Rosemblat, Cognitive Science Branch, NLM
Dr. Angela Ruffin, Office of the Director, NLM
Mr. Jerry Sheehan, Office of the Director, NLM
Mr. Mark Siegal, Division of Extramural Programs, NLM
Dr. Elliot Siegel, Office of Health Information Program Development, NLM
Dr. Hua-Chuan Sim, Division of Extramural Programs, NLM
Dr. Charles Sneiderman, High Performance Computing & Communication, NLM
Ms. Marti Szczur, Division of Specialized Information Services, NLM
Ms. Michelle Trout, Office of Federal Advisory Committee Policy, OD
Ms. Emily Vardell, Associate Program, NLM
Dr. Deborah Zarin, Lister Hill Center, NLM
Dr. Frederick Wood, Office of Health Information Program Development, NLM
Dr. Jane Ye, Division of Extramural Programs, NLM

I. OPENING REMARKS

Dr. Cynthia Morton, Chair of the NLM Board of Regents, welcomed the Regents, alternates, consultants, and guests to the 146th meeting of the Board. She introduced three new Regents: Dr. John Connolly, Dr. Carol Friedman, Dr. Louis Rossiter, and Bruce James. (Mr. James was unable to attend). She also welcomed Gail Graham, an alternate ex officio member from the Veterans Administration, Ms. Kathryn Mendenhall an alternate ex officio member from the Library of Congress, and Dr. Holly Buchanan of the University of New Mexico, a consultant.

II. REPORT FROM THE OFFICE OF THE SURGEON GENERAL, PHS

Rear Admiral Helena Mishoe, representing Dr. Kenneth Moritsugu, Acting Surgeon General of the U.S. Public Health Service, reported that Dr. Moritsugu was retiring on September 30, 2007 after 37 years of service to the Commissioned Corps. A new Acting Surgeon General should be in place soon thereafter. She added that he apologized for not personally attending the Board of Regents meeting due to a press conference scheduled for the same time. The Office of the Surgeon General has been extremely busy in the area of science and communication. On September 12, 2007 the proceedings of the Surgeon General's Deep Vein Thrombosis Workshop were posted to the Surgeon General's Web site: www.surgeongeneral.gov. Also, the proceedings of the Surgeon General's 2006 workshop on Health Care Literacy should soon be posted on the Web site as well. Two other workshops that are expected include "Making Prevention of Child Maltreatment a National Priority," and the Surgeon General's workshop on "Women's Mental Health." Admiral Mishoe also discussed the Prematurity Research Expansion and Education for Mothers who deliver Infants Early Act, or "PREEMIE Act," designed to educate the public about the high rate of premature births in the United States. It requires the Surgeon General to hold a workshop, which should be in early spring. Dr. Moritsugu missed

today's board meeting because he was at a press conference with representatives of the EPA and CDC to discuss the health consequences of involuntary exposure to tobacco smoke. Another priority of the Surgeon General's Office is also a priority for Secretary Leavitt — personalized health care. Mishoe thanked Betsy Humphreys, Clem McDonald and others at NLM for their help on this effort. Lastly, she pointed out that Acting Surgeon General Moritsugu is actively engaged in promoting the issues raised in the Surgeon General's Call to Action to Prevent and Reduce Underage Drinking, or PUD. He is traveling across the country discussing how underage drinking should not be considered a right of passage for underage teenagers. Dr. Donald Lindberg noted that Dr. Moritsugu has been a wonderful Acting Surgeon General and friend of the National Library of Medicine. Dr. Morton agreed and said she took Dr. Moritsugu's message on underage drinking to her son's school, where teachers have involved students in crafting solutions to the problem.

III. MAPPING COMPLEX TRAITS IN THE DOMESTIC DOG

Dr. Elaine Ostrander, Chief and Senior Investigator of the National Human Genome Research Institute (NHGRI) Cancer Genetics Branch, reported on the work she is doing on mapping complex traits in the domestic dog. Her laboratory is currently developing resources to identify genes commonly associated with complex traits in the dog. The traits include cancer, the number one killer in dogs. Dr. Ostrander reported that half of her lab works on human genetics, breast and prostrate cancer as well as bladder. The other half has taken the lead in developing the canine system as a way of understanding genes in human health and biology. She reported on two recent papers that have come out of her lab, one featured on the cover of *Science* and the other on the cover of *PLoS Genetics*. There are 157 breeds of dogs recognized in the United States and every breed is fixed for a certain set of traits, body size, color, or leg length, for example. Researchers have postulated that a small number of genes are likely responsible for the complex traits that we are interested in studying. Dr. Ostrander outlined their approach for studying body size. Initially, her lab studied the Portuguese Water Dog (PWD), to find the regions of the genome that they thought would be important for controlling body size. They analyzed a large sample of PWDs to narrow the interval of genes and took a lot of samples from many dogs, big and small, to narrow the problem down to a single nucleotide of a gene. This is now used as the paradigm for all complex traits that they are studying in dogs. There are about 10,000 registered PWDs in the U.S. and they all descend from 30 dogs that came here 50 years ago from Portugal. They started a project at the University of Utah called the "Georgie" project, where about 1000 dogs are enrolled. They collected DNA and interrogated their genome from 500 different positions using "markers." Each of the dogs gave live X-rays and from them they were able to collect 91 different measurements. They used "principal component" (PC) analysis to classify these measurements, or phenotypes, into systems of correlated traits. They were interested in the striking observation that there is a 40 fold difference in size between the very largest dog, the St. Bernard, and the very small dogs, the Chihuahuas, etc. They took 42 measurements, analyzed them, and found, as expected, that body size is inherited. They found seven places in the genome where they believe that there are genes that affect a dog's overall

body size. Chromosome 15 is responsible for body size. They identified the region of the gene that makes big dogs big and little dogs little. Their findings were featured on the cover of *Science* Magazine. They looked at other types of DNA samples representing all breeds in the US and looked at where differences were the greatest. They collected samples from big and little dogs — collecting over 10,000 DNA samples. They looked at where the genetic variation was the greatest in the genome between 10 giant breeds and 10 very small breeds — and again, it was in the same chromosome. They looked at where variation changes take place genetically and were able to go from a whole genome, to a chromosome, to a 15 million base pair region, to a single gene, and then down to a small number of bases within a gene. It was a nice study. They are now examining other traits like leg length, face types and other unusual canine features, and they are also interested in differences in dog performance. They decided to look at over-muscling in Bully Whippets — yet, the Bully Whippet phenotype is qualitatively similar to other animals. There is a mutation that is responsible for the muscling and improved racing performance. This finding has interesting implications for humans. Are mutations found in other high performance groups? Dr. Ostrander said that this is the number one question they get.

Dr. Lindberg commended Dr. Ostrander for her presentation and asked her to share her story about the goldens and the wolves. Dr. Ostrander said that the closest relative to the dog is certainly the wolf. They have studied wolves all around the world trying to find which wolf evolved into the dog and they have concluded that the dog evolved from a wolf in Eastern Asia. Dogs were probably domesticated multiple times with multiple events taking place to encourage domestication. Goldens are a completely contrived breed formed by crossing many different types of dog breeds and is now one of the most popular breeds of dogs. Dr. Walker pointed out that different dogs have different types of diseases. He asked if she was looking into that. Dr. Ostrander said they have a lot of studies ongoing in that area. For example, bladder cancer is an important study in dogs because only about five breeds get bladder cancer — Scottish terriers are thirty-fold higher than other breeds. Dr. Friedman asked about a mouse discussed in Dr. Ostrander's talk that had a knockout of IGF1, and she asked if that information is used cross-species? Dr. Ostrander said that comparative genomics is. Dr. Morton observed that Dr. Ostrander may be able to look at the effects of natural selection. Dr. Ostrander said they are really interested in the genetics of domestication. In other words, what genes were acted on to make the transition between the wild animal to being man's best friend. They are collecting DNA samples and it has been challenging. Dr. Morton said the gene responsible for human height would be of interest. Dr. Ostrander agreed.

IV. CONSIDERATION OF MINUTES FROM PREVIOUS MEETING

The Regents approved without change the minutes from the May 8-9, 2007 meeting.

V. DATES FOR FUTURE BOARD OF REGENTS MEETINGS

The Board of Regents will meet next on February 12-13, 2008. The Board meeting next spring is

May 13-14, 2008. The dates of September 16-17, 2008, were adopted for the following meeting.

VI. REPORT FROM THE NLM DIRECTOR

Dr. Donald Lindberg reported that Dr. Jordan Cohen could not be present because he is chairing a meeting in Qatar but will be present at the next meeting. With respect to NLM's budget, Dr. Lindberg noted that the House passed the Appropriations bill and the Senate increased it slightly. Both reports were included in the Member's packets.

In the case of personnel, Dr. Lindberg noted that Lisa Lang is the new Assistant Director for Health Services Research Information. She comes from the Centers for Medicare and Medicaid. She had recently suffered a fall and was unable to attend the Board's meeting. He also introduced the NLM's new Associate Fellows: Grace Ajuwon of Nigeria is participating in the program as an International Fellow who received her degree from the University of Ibadan and is a reference and information services librarian in the E. Latunde Odeku Medical Libra; Sarena Burgess (University of Tennessee, MLS 2007); Kate Flewelling (University of Wisconsin, Milwaukee, MLS 2007); Joanna Karpinski (Drexel University, MLS 2007); Brenda Linares (UCLA, MLS 2007); Melissa Resnick (Queens College, MLS 2007) and biopsychology also from the Rensselaer Polytechnic Institute); Emily Vardell (Texas Woman's University, MLS 2007,also a former Fulbright scholar.)

Dr. Lindberg noted a number of retirements and deaths at NLM. Bob Mehnert, who headed up NLM's Office of Communication, retired after more than 43 years of service. He said that Bob would be missed as he was the skillful producer of gracious prose, and possessed exceptional editorial and literary talents. Karen Hajarian retired after 15 years of service. She had performed a great job managing NLM's efforts to expand usage of its databases. Dr. Bill Hole, a physician who worked very closely in the development of the Unified Medical Language Systems (UMLS) Metathesaurus at NLM, also retired. Lastly, he pointed out that Vera Hudson, who worked on NLM's Hazardous Substances Database and many other important initiatives within SIS, passed away and will be much missed.

Extramural Programs Director Dr. Milton Corn introduced new staffer Dr. Jane Ye, who is heading NLM's bioinformatics grants program. She earned her bachelor's degree in molecular genetics from Cornell and her doctorate in molecular biology from Dartmouth. She previously worked at the NHLBI, where she managed a large portfolio of genomic projects.

Dr. Lindberg briefly covered legislation and noted that a bill is under consideration that would impact clinical trials. ClinicalTrials.gov was begun in 1999 following the FDA Modernization Act of 1997 which ordered the NIH and the NLM to create a registry of trials. That registry has grown from zero to over 43,000 and has been a big success. Now we are facing the problems of success, and one of those is that the Congress would like to have greater transparency about the results of clinical trials, including statistical information and narrative summaries for technical

and non-technical audiences. He indicated that Dr. Deborah Zarin would discuss issues surrounding this legislation in her presentation. With respect to Health IT, Dr. Lindberg said that information regarding pending legislation is included in the Members' packets. In the area of genetics and genomics, Dr. Lindberg said that the Director of the Genomics Institute, Dr. Collins, has been outspoken for years on the need for legislation to prevent discrimination based on genetic information such as denial of insurance. A bill has been approved by the House, although not yet adopted by the Senate, to protect the public from the misuse of genetic information.

Acting on a recommendation of the 2006 Long-Range Plan, Dr. Lindberg noted that he has created a NLM-wide Disaster Information Management Steering Committee made up of members from each division (Tab D). He wants to explore a unified approach, building upon the use of Go Local in improving access to community-specific disaster mitigation information. Experiments are already underway in Missouri and North Carolina and in 23 other states. Dr. Lindberg went to the opening of Go Local in Minnesota. Betsy represented NLM at the launch of Go Local in Illinois soon thereafter. Both of these recently launched projects, like the others that are in place, are unique in their approach to connecting their communities with health care information. NLM will look at the Go Local projects as the Disaster Information Management Research effort moves forward. Dr. Steven Phillips, former NLM Deputy Director for Research and Education, has been tapped to head up this project.

Under Tab E, Dr. Lindberg discussed NLM's collaboration with the National Agricultural Library (NAL), the American Veterinary Medical Association (AVMA), the US Agricultural Information Network (USAIN), and the Veterinary Medical Libraries Section (VMLS) of the MLA to create an information portal to improve information services for clinical veterinarians. In response to a question from Dr. Walker about the percentage of the veterinarian literature available from NLM databases, Sheldon Kotzin replied that about 97 journals were in the system and that PubMed has become the primary source for veterinarians engaged in research, but not for the clinical veterinarians. This concerns Dr. Lindberg and that is why the new portal is in production.

Dr. Zerhouni asked Dr. Lindberg to chair a Trans-NIH Biomedical Informatics Coordinating Committee -- a committee on clinical informatics of which clinical informatics would be a subset. Tab F gives the names of the members from the Institutes and Centers. They have had one meeting. The Committee is to coordinate and inform everyone at NIH about what is going on and also to represent Dr. Zerhouni at other government meetings. Next, Dr. Lindberg described how the NIHSeniorHealth Web site has been a real success. NLM worked with NIA to develop this site, and now a number of other Institutes are collaborating to add new health topics formatted with senior-friendly features, including a "talking" function. Television and radio public service announcements (PSAs) will begin to air this fall, and Dr. Lindberg presented one of the PSAs. Dr. Walker asked Dr. Lindberg if NIHSeniorHealth is produced by the NLM. Dr. Lindberg responded that NLM works with NIA to produce the site. Dr. Lindberg observed

that many of the features are good for young and old alike. Dr. Rossiter was asked if the Centers for Medicare & Medicaid Services (CMS) could help get the word out about NIHSeniorHealth. Dr. Lindberg thought that was a good suggestion.

Dr. Lindberg noted that NLM continues to produce wonderful exhibitions and then when possible, travels them. The *Frankenstein* show is such an example. It traveled to over 80 cities. *Changing the Face of Medicine* is traveling to 60 cities nationwide. A smaller show called *Opening Doors: Contemporary African American Academic Surgeons* was successfully exhibited here at the NLM and its tour schedule is in the Board book.

Lastly, Dr. Lindberg talked about the Annual Informatics Training Conference held at Stanford University on June 26-27, 2007. Next year, the Conference will be held here at NLM. Dr. Milton Corn mentioned that we invite CDC public health trainees and also ask the Veterans Administration, which does some informatics training, to send people as well. Dr. Corn noted that NLM is probably the largest source of informatics training in the country and he believes the conference provides a good networking opportunity. Dr. Lindberg commended Dr. Corn and Dr. Valerie Florance for the great job they have done in organizing it.

Betsy Humphreys introduced Paula Kitendaugh, a former NLM Associate who has been named head of the Reference and Web Services section.

Dr. Walker pointed out that no mention had been made of the new search engine, Vivisimo. Dr. Lindberg said that a full report and demonstration will be given at a future meeting. He noted that NLM selected Vivisimo as its new search engine instead of Google because the latter didn't seem as responsive to NLM's specific interest in being able to customize the software.

Asked about NIH's public access policy, Dr. Lindberg reported that in the House- and Senate-passed versions of the Appropriations bill there is a provision that would make public access mandatory. There would be a twelve month delay before manuscripts would have to be deposited in PubMed. If the bills go through with the public access provision, implementation will be discussed at our next meeting.

VII. CLINICAL TRIALS INFORMATION

Dr. Deborah Zarin discussed how ClinicalTrials.gov is a prospective Web site and registers both interventional and observational trials before they have begun The database links to publications reporting results, but does not have unpublished results. ClinicalTrials.gov was created in response to the Food and Drug Modernization Act (FDAMA) to provide the public with information about clinical trials. The International Committee of Medical Journal Editors (ICMJE requires registration of trials as a prerequisite for consideration of publication of their results, and this has had a major impact on the expansion of the database. There is also a law in Maine that requires registration of trials in ClinicalTrials.gov. About 19 other states are

considering similar laws. It was AAMC that convinced Maine not to adopt is own trials registry but to rather mandate registration in the ClinicalTrials.gov database. There are a number of Attorney General settlements requiring registration in ClinicalTrials.gov. And the WHO has also promulgated policies around registration in ClinicalTrials.gov as well.

Dr. Zarin said that we are getting about 200 new trials each week. NLM has nearly 45,000 trials now of which 85% are interventional trials with drugs. There are about 2,000 device trials. Trials come from all sponsors: NIH sponsors about one-third of the trials, and industry just below 30%. NLM will have a new search engine in place in coming weeks which will improve display of trials. A key issue for NLM is how to ensure that users will be able to find a complete non-redundant list of trials that meet a certain search criteria. Improved search engine functions will improve our ability to do this. A problem involves the use of serial numbers rather than drug names to identify the drugs involved in trials. Nobody, including the FDA, keeps complete track of serial numbers and how they map to subsequent drug names, which makes it challenging to provide complete retrieval of the trials involving particular drugs. The percentage of trials that use serial numbers, by phase, are: 43% in phase 1, 35% in phase 2, 11% in phase 3, and 1% in phase 4. There are also a lot of problems with devices. They have multiple components. They change incrementally but keep the same name. Trial duplication is also a problem. A trial can be registered more than once. There are many multi-site trials. About a third of our trials are located in more than one site. About half of those are only in the United States. We have one trial in 46 countries. Duplicate International registration is problematic. NLM is working with WHO and the ICMJE to attempt to minimize it.

Dr. Zarin pointed out that the pending FDA legislation contains a provision that expands registration requirements with sanctions for noncompliance. The only legal mandate in this country right now is to register trials of drugs for serious or life-threatening conditions, although ClinicalTrials.gov receives many other trials. There is an increasing interest in expanding the scope of registered trials and also in results reporting. Both Houses have passed bills that are currently in conference, and action is expected soon. Both bills talk about device trials and there is a possibility of a "lock box" which is a term for trials that have to be registered but not posted publically until some trigger event occurs. Both give NIH new mandates that they will have to meet with respect to the trials they fund. Another provision of this new legislation involves the linkage of clinical trial data to FDA data, and there will be issues involved in meeting this new objective. There are also concerns about a new results database. Evidence-based medicine involves more than just the results of the clinical trial. It also involves consideration of the trial design. You need to see the whole body of evidence and NLM needs to think about this to design a results database that will be meaningful. The individual trial is often the wrong unit of analysis. Nevertheless, there is a desire to enhance transparency. A trans-NIH work group is in place to determine how to get NIH-funded studies out to the public. We are also looking at existing results databases in industry to see what has worked well and what we can do better.

Dr. Zarin said the expected legislation will require the new results database to be in place within

a year. Both passed bills discuss a "narrative lay summary." They want the data providers, like Merck or Harvard, to provide a summary of the trial. So even though the data are coming from the sponsor, the consumer might think the information is coming from the NIH. Issues involved include: How do we facilitate rational use of the data once it is up? Nobody knows how to do this. NLM will be charting new territory. Other questions are: Who is the intended audience and how are we going to evaluate the results databases? Other issues to ponder: What constitutes a "minimally acceptable" results record? This is a complicated issue. And what about disputes that arise about validity? How should NLM handle calls about a trial's outcome when NLM does not have the actual data? To what external sources should we link?

Dr. Morton mentioned the possibility of establishing a working group on clinical trials reporting. She read a proposal for consideration of the Board detailing how an NLM Board of Regents Working Group on Clinical Trials Reporting would meet and report on its findings. She read the following "Proposed Charge for Working Group on Clinical Trials Reporting:"

> The NLM Board of Regents Working Group on Clinical Trials Reporting will advise the NLM Board of Regents on issues associated with the expansion of Clinicaltrials.gov and the addition of a results database. It will take into account relevant legislation and consult as necessary with relevant stakeholders and potential users of the system to provide guidance on initial implementation issues and longer-term strategies for enhancing: data content and format; procedures for submitting, storing, and posting information; quality assurance; the provision of summary information for the general public; the scope of clinical studies included in the databases. The Working Group will meet at least twice per year and prepare a report, with any recommendations, for presentation to the NLM Board of Regents. [September 18, 2007]

Dr. Morton asked if there was any discussion about the proposed working group. Dr. Detre said that during the first year of implementation, perhaps more than two meetings would be indicated. Betsy Humphreys said that the phrase "at least twice a year" would provide the flexibility to accommodate more meetings, if needed. Dr. Lindberg expressed support for the establishment of such a working group and noted that the task will be to let the world know what NLM is doing to try to obey the law. He said this is another classic unfunded mandate. NLM is mentioned in the legislation and is not trying to dodge its responsibility, but would like to approach it more gradually than do it in a year. For example, when they say, "Inform the public" — what does that really mean? How should the public be informed? That question itself is probably a researchable question. Dr. Rossiter asked Dr. Lindberg if the conferees were aware of this concern. In response, Dr. Lindberg said that we are trying to make them aware of our concerns. Rossiter also asked if the notion of an experimental period or a demonstration for this had been considered. Dr. Zarin mentioned that the Senate bill had called for a feasibility study and the house bill says do it within a year. Dr. Lindberg responded that NLM, therefore, is not just sitting around doing nothing. Dr. Detre asked about the patient who looks at the

ClinicalTrials.gov Web site and decides to ask his doctor about the trial. If the doctor doesn't know about the trial, once the question is asked, shouldn't there be something in the legislation that requires academic institutions or teaching hospitals to provide further guidance? Dr. Lindberg said he didn't believe that is covered by the legislation. Dr. Detre said that the issue is whether or not there should be a network of teaching hospitals and medical centers to provide information in certain subspecialty areas. Dr. Rossiter commented on the fact that the NLM is simply posting information that others give us. He asked if we know who is behind the legislation – well-intentioned staffers on Capitol Hill? Dr. Lindberg said he believes it is the principals. Dr. Zarin pointed out that many people have been behind the legislation. She said that most agree that it is a good idea, but some don't believe that the public should be the intended audience. Dr. Lindberg said that Congressman Waxman and Senator Kennedy were instrumental in the bill's progress and that conferees had not yet been chosen. Dr. Rossiter asked if there is any possibility of interpreting the legislative language so that NLM is merely creating the field, not completing it? Then it would be clear that NLM is merely making the information available, cataloguing it. Dr. Zarin said that even if it is stated, when it is under an NIH banner, some people will assume that it is endorsed by the NIH. Dr. Rossiter asked if clinical guidelines have warnings (National Guideline Clearinghouse) and that might be done for ClinicalTrials.gov too. Dr. Connolly noted that FDA has committees, like device committees, and couldn't we get some advice from them about how to set this up? Dr. Lindberg responded that FDA has the responsibility as well and the question is how we work with them. When FDA approves a drug for sale, then they should open the file and show the stuff they have and then we could link to it. That usually doesn't happen now. Dr. Zarin remarked that that was one of the IOM's drug safety recommendations. Dr. Connolly asked if the Framingham study was a clinical trial and Dr. Lindberg said it was observational and noted that much of what we know about cardiovascular disease came out of that study and now they are adding a genetic component to that. Dr. Detre recommended that Dr. Zarin's presentation be offered for publication in the *JAMA.* Dr. Zarin said that it had been published several months ago. Do we ever talk about procedures being added? Dr. Zarin mentioned that about 10,000 trials have procedures as one of the interventions. The current law is drugs only and the proposed law will include what FDA regulates: drugs and devices.

The Board endorsed the establishment of the Working Group, which Dr. Morton will chair. Dr. Rossiter also volunteered to participate. Dr. Lindberg indicated that Dr. Cohen also said he would serve on such a group. Dr. Morton also asked the Board of Regents to approve in concept an NLM initiative to study the problem: how to present summary information to the public concerning potentially complex scientific studies and their potential relevance to specific patients. Ms. Stanley made a motion that the BOR initiate the study per the language in the BOR book. The motion was unanimously approved.

VIII. PRESENTATION OF REGENTS' AWARD

Dr. Morton presented the NLM Regents' Award for scholarship or technical achievement to Dr. William Hole. He recently retired after 18 years at the NLM after heading up the development

of the UMLS Metathesaurus. His award reads: "For outstanding leadership in the design and development of successive generations of the UMLS Metathesaurus, a major force in advancing biomedical informatics research and the state of art in medical terminology."

IX. SEMANTIC MEDLINE: ADVANCED TECHNOLOGIES FOR BIOMEDICAL INFORMATION MANAGMENT

Dr. Thomas Rindflesch discussed the reason for the research resulting in Semantic MEDLINE. NLM provides a major source of access to biomedical research, namely PubMed/MEDLINE. A researcher wanting to get the most recent research on breast cancer, for example, will retrieve many articles. The challenge is to identify the new and interesting content in the vast and increasingly complex scientific literature. The new tools NLM is creating are designed to address complexity and manipulate information to help us navigate the literature. "Semantic MEDLINE" is one such application being developed by the Lister Hill Center. Its core characteristic is that in the sea of text that you get back from a MEDLINE/PubMed search – Semantic MEDLINE identifies salient information and converts it into relationships in computable form, which are visualized as an informative graph with links to the original citations. The ability to connect salient information across documents and structured knowledge sources helps users keep up with the research literature and discover connections which might otherwise go unnoticed. Dr. Rindflesch noted that this new process would not be possible without NLM's Metathesaurus. He gave a live demonstration of Semantic MEDLINE, looking at the most recent 500 citations on breast cancer and the graph of relationships it detects within them. which provides a pretty informative view of what is known about breast cancer, including treatment. He presented another example related to colon cancer research and then showed how the use of Semantic MEDLINE helped to sort through relationships between research articles to collect information pertinent to scientific inquiry. NLM is collaborating with the NHLBI to use this technology in clinical practice guideline development. NLM hopes to make this process more efficient and effective.

Dr. Rindflesch said preprocessing relationships was the most time-consuming aspect of the project. When asked about problems involved in processing relationships in multiple languages, Dr. Rindflesch said the Metathesaurus has many non-English vocabularies that amount to synonyms — so NLM can output information in multiple languages. Social science concepts are particularly challenging because the language used to express them is less precise. Dr. Walker asked how many articles in PubMed can be retrieved in free full text. Betsy Humphreys said about 12% overall and 18% for more recent articles, noting that the number is creeping up. Dr. Harris asked if Dr. Rindflesch envisioned a time when this kind of tool could be available for use with clinical trials. Dr. Rindflesch said that he believed that it could be used in clinical trials and could be made available to the public if an effort was made to edit the output to avoid errors. Dr. Rindflesch said that once finished, the final version would allow the user to manipulate the data any way they want. Ms. Stanley said that the concept of visualizing relationships is being used in other research areas as well, like the extension of the UMLS to a consumer vocabulary using

terms already familiar to consumers. Dr. Rindflesch said that they are working on the development of a consumer vocabulary in conjunction with researchers at Brigham and Women's Hospital. Dr. Lindberg observed that there have been about a dozen ventures that made approaches to respond to the consumer vocabulary issue and mentioned that he didn't know whether it was essential or not. But if you look at the questions that people bring to MedlinePlus, you will find that 85 percent of the people know either diagnosis or name of a drug. It means that they have some sort of medical encounter and left without all of the information. So at least you would want to handle those variations. Ms. Stanley agreed and noted that if they have left the medical encounter without all the information they need, what is the question they will be asking? Perhaps NLM can prepopulate some of those questions and answers and keep the information up-to-date. Dr. Lindberg mentioned he just returned from Utah where someone said to him, "Why don't you just put an asterisk next to the good articles, and take the rest of them out?"

X. EXTRAMURAL PROGRAMS REPORT

Extramural Programs Director Dr. Milton Corn presented a snapshot of NLM's R01 research grant program, in the larger context of what's happened at NIH over the last four or five years. Between 1999 and 2003, Congress doubled the NIH budget, providing annual increases of about 15% annually, compared to the past years' rates of 4-5%. This development held great promise for advances in biology and, in turn, breakthroughs in diagnosis and treatment. Certainly, there were many positive outcomes, but also some surprises. The number of R01 research grants (the "workhorse" grant at NIH) did not increase significantly. The success rate, that is, the chance of receiving a research grant, did not increase but actually decreased. And, the number of published papers by U.S. scientists did not increase significantly following the doubling. Against this backdrop, Dr. Corn described the grants situation at NLM. Funding for NLM's R01 grants doubled between FY96 and FY03, and actually almost tripled, in absolute dollars, from FY95 to present. Despite that fact, NLM did not fund a large number of new awards and, surprisingly, the success rate fell from 23% in FY96 to 25% in FY01, and to10% in FY06.

One reason for the plummeting success rate was a jump in the number of applications. In FY96, NLM received 60 R01 proposals, but in FY06, perhaps because people knew that more money was available, 124 applications came in. Why were there no more grants awarded, during a funding upturn? The average cost per R01 award went from $206,000 in FY96 to $413,000 in FY06, and, in turn, the same number of proposals received support *after* the doubling as before. Dr. Walker asked why the amount of the average award increased. Dr. Corn replied that people applying knew that the pot of NIH money had increased dramatically. In addition, a restriction imposed by the NIH grants office, requiring extra paperwork for proposals amounts over $500,000, spurred applicants to float their application amounts higher, into the $400,000's.

Dr. Walker asked whether NLM should be disturbed that the number of scientific papers didn't increase. Dr. Morton said that the figure might be misleading, because sometimes there are large

consortia working on "big science" projects like genome-wide association studies. Ms. Stanley asked whether NLM should be concerned about the many proposals it *isn't* funding. Dr. Corn said that NLM is a microcosm of NIH, and that additional funding will of course ensure that more worthy projects are funded. For the foreseeable future, though, increased budgets seem unlikely. Finally, Dr. Corn remarked on a dramatic shift in the kinds of projects NLM has funded in the past decade, reflecting the rise in the use of computing. Around 1990, all grants awarded were clinical in nature. In FY96, 45% of NLM grants were clinical in nature, and 13% related to basic biology. The clinical/basic rates were 57% and 13% respectively among FY06 grants at NLM. As NLM grants are increasingly centering on computational biology, other ICs are also getting involved in this area. The National Human Genome Research Institute (NHGRI) and the National Institute of Mental Health (NIMH), were among the first to award funding in that area, and today, many of the informatics topics that NLM alone used to support are getting money from the "bigger fish" Institutes and Centers, like the National Cancer Institute (NCI).

Dr. Corn discussed some potential Research Grant Initiatives for FY 2008. To increase the impact of NLM grant funds, and to focus some resources on topics of particular interest to NLM and the Long Range Plan, EP plans to issue one or more Requests for Applications (RFA) in relevant areas in FY 2008. Such RFAs will have the format of challenge grants with emphasis on innovation and with the use of abbreviated application processes.

Subjects under discussion for RFAs include: data-mining for *in silico* discovery; "intelligent" retrieval; obtaining data from heterogeneous databases; self-managing knowledge acquisition for databases; automated simplification of scientific information for consumer use; archiving of clinical data; automated annotation; disaster information management tools; virtual universes (e.g. video games) for behavioral studies and for education; personal health record informatics; automated summary of patient's clinical status; human-computer interface issues, including visualization of data; and Natural Language Processing, with an emphasis on clinical free text. Discussion touched on relative merits of a number of the options, although no specific preference was defined. Dr. Lindberg suggested that modeling be added to the list of contenders. Attendees also spoke about the need for budget increases, the effects on the research community of low success rates, and the implications of steadily rising research costs.

MEETING CLOSED FOR REVIEW OF GRANT APPLICATIONS - 2:40-3:00 P.M.

XI. MEDLINEPLUS EN ESPAÑOL OUTREACH AND PROMOTION

Ms. Paula Kitendaugh, head of NLM's Reference and Web Services Section, described the evolution of MedlinePlus en español, its target audience and the lessons that her office has learned since the site's creation in 2002. MedlinePlus in English came first, in 1998, and MedlinePlus en español (henceforth "M+ esp" in these minutes) followed in September 2002. M+ esp came about because, between 1998 and 2002, MedlinePlus users cited Spanish language content as their number one need. There are 40 million Hispanic Americans (U.S. Census,

2000), three-fourths of whom (28 million) speak Spanish in the home. Census data also show that a subset of Spanish-speaking Americans, about 14 million, self-identify as speaking English "less than very well." MedlinePlus staff also heard from health care providers, hospital and public librarians, and others who had patients or clients needing health information in Spanish, so that they could read it with ease. Generally, Hispanics don't have Internet access at home, but many do at work and at school, and, most of all, at public libraries.

One unexpected finding in NLM-conducted focus groups was a strong preference for English language information, which Hispanics perceive as being of higher quality. They especially rely on English language sites for medical information, perhaps because it is more U.S.-centric and relates to the American health care system, treatment protocols, etc. Focus group participants noted that online information in English was much more detailed and complete. They noted, too, that the poor translation quality and large number of spelling and grammar errors made online health information in Spanish appear untrustworthy.

Within the first year of its launch, M+ esp was receiving one-third of all MedlinePlus traffic. That trend continues today. In 2006, 26 million unique visitors viewed 207 million pages on the site. MedlinePlus, in both English and Spanish, has consistently been among the top five federal sites among Internet users surveyed by the American Customer Satisfaction Index. MedlinePlus en español also has a global reach. There are 322 million Spanish speakers worldwide, and many are finding their way to the NLM site. Google/Mexico and Google/Spain direct many searchers to MedlinePlus en español: 42% of the site's traffic comes from these countries. Many users of M+ esp, here and abroad, are medical students, health care providers, researchers and other professionals who hold NLM and NIH in high regard. Despite that encouraging news, M+ esp is still not reaching much of its U.S. audience. Ms. Kitendaugh showed a U.S. map illustrating a wide gap between the number of Spanish speakers per state and the number who use M+ esp. Data show that people who discover M+ esp resoundingly approve of it. How, then, to help them to find it in the first place? NLM has begun a one-man outreach plan, in the person of popular TV personality Don Francisco, sometimes called the Oprah of the Spanish-speaking world. Don Francisco, whose Miami-taped TV program, "Sabado Gigante," attracts some 100 million viewers in over 20 countries, also has sterling credentials as a health advocate for the Hispanic community. Public service announcements he made to promote M+ esp were distributed to 100 Spanish-language TV and 450 Spanish-language radio stations and aired nearly 30,000 times in the first six weeks. M+ esp saw marked usage increases in several markets after the PSAs aired, and glowing reports came in from managers and programmers at participating radio and TV stations. In closing, Ms. Kitendaugh mentioned that radio and TV, and a popular celebrity, has a place in reaching out to the Hispanic audience, but these constitute only a first step. As in any public affairs campaign, multiple media and channels are needed, and the MedlinePlus team will continue to explore these, in its attempt to broaden use of the site among Spanish speakers in the U.S. and abroad.

Dr. Rob Logan of NLM's Lister Hill Center introduced Mr. Jorge Lambrinos, Director of the

Edward R. Roybal Institute for Applied Gerontology at the University of Southern California, Los Angeles. The Institute is under contract to NLM to measure the accessibility and acceptability of MedlinePlus en español and its related educational materials. Dr. Logan noted that the Roybal Institute works closely with community clinics and health care institutions in southern California, and carefully follows the ways in which Latinos obtain health information.

The NLM-funded study has three elements: (1) evaluation of the M+ esp Web site as a valuable education tool for the Latino population, (2) development of a media campaign to reach physicians and consumers, and (3) a study to determine whether M+ esp improves the health outcomes of its users. To ascertain whether M+ esp was (or could be) a valuable, empowering source of health information, the Roybal Center assembled seven focus groups of diverse age groups from young adults to the elderly. These groups included consumers, physicians, health advocates and family caregivers, as well as community health professionals and social service providers. Among the findings: the Latino young adults, age 25 and under, tended to trust sites ending in .gov or .org more than those ending in .com, and were much more familiar with MEDLINE than with MedlinePlus; the Latino caregivers and care advocates (age 35-60) were fairly literate about health issues, but tended to go to the library and consult printed material like magazines, rather than searching the Web. None had heard of M+ esp; the Latino older adults (age 60-85) bore out a truth that seemed to run through all groups — that the women in the family were much more concerned about health than were the men. In this group, most relied on health information from their HMOs, local clinics, health fairs and printed sources. Two had been on the Internet and had consulted WebMD. The monolingual (Spanish-speaking), foreign-born population looked to Spanish-language radio and TV stations for health information, and some also reported using the computers at their local library. Among the health and social services providers, half had heard of MedlinePlus, but were not aware of the Spanish language version. Among the Spanish as first language group, members of whom prefer speaking Spanish at home, most used the Internet but few relied on it as a primary source of health information. The focus groups yielded these findings: (1) All groups had a strong positive response to M+ esp, and felt that if they knew about it, Latinos would definitely use it; (2) Trusted media personalities, such as Don Francisco, are valuable for spreading the word, as are trusted grassroots organizations; (3) Challenges included misinformation in the community ("old wives' tales") and a sense of fatalism — that illness is God's will and people have no control over it; (4) "Promotores," individuals trained to go into communities and distribute health information, could be excellent M+ esp ambassadors. (The Public Health Service uses them frequently to distribute information to the Spanish-speaking community.) Mr. Lambrinos recommended the promotion of M+ esp among doctors and patients, partnering with National Community Health Centers, among other organizations. The Spanish-language promotional CD should be distributed to doctors, to be played in their waiting rooms. M+ esp materials should also be sent to grocery stores, pharmacies and other places where Latinos go to pick up prescriptions. Federal agencies like the Administration on Aging and the Centers for Disease Control and Prevention should be encouraged to link to and actively promote M+ esp, and grassroots organizations could certainly do the same. Mr. Lambrinos remarked that MedlinePlus, in

English and in Spanish, is a great source of reliable information, and efforts should be made to expand its use in underserved segments of the population, among Spanish-speakers and others not yet aware of its riches.

Mr. Chabran mentioned that a recent Pew Internet study, *Latinos Online*, focused on Spanish-dominant households in the U.S. Among that group, only 32% go online. Among foreign-born Spanish speakers, only 43% are Internet users. Factors influencing those figures were education, income and English language proficiency. He said that, above all, NLM should work to make sure Spanish speakers know about Web resources like MedlinePlus en español. The National Network could help in this regard. Ms. Frierson from the National Agricultural Library suggested NLM explore linkages with *Prevention* magazine, the Red Cross, AARP, the Agency for International Development and CNN International, which runs PSAs in many languages. Dr. Lindberg asked whether Sirius satellite radio should be considered as another publicity outlet. He added that he was aware of the importance of promotores in Texas, where the Library has collaborated with them, but asked Mr. Lambrinos if they were a nationwide network. Mr. Lambrinos said that they are very active in California and have expanded to include many states. They are very trusted among the Latino community, which is crucial because, if people trust you, they'll trust the information you present. Mr. Chabran concurred, but said that the promotores work in a very non-digital, low-tech medium. Perhaps they could use cell phones or PDAs to spread the word about M+ esp. Because many clinics are financially strapped, Mr. Lambrinos suggested NLM attach some money to its M+ esp promotional efforts at those sites. Dr. Friedman suggested NLM educate schoolchildren about M+ esp at an early age, so they will use it to glean health information for themselves and for older family members. Mr. Chabran said that Podcasts and other multi-media promotion would be critical to reaching younger Latinos. Ms. Stanley asked whether the Don Francisco PSAs are on YouTube; yes, Ms. Kitendaugh replied. Ms. Stanley asked whether a younger personality, such as Colombian-born singer Shakira, might join Don Francisco as a M+ esp spokesperson. She also recommended that Go Local appear in Spanish, to help Latinos find health care resources in their neighborhoods. Mr. Lambrinos said that it would be helpful if the PSA content could be adapted to health magazines, women's magazines and other print organs. If the TV PSAs appeared during the popular Spanish language soap operas, "novellas," they would reach a wider audience, too. Ms. Stanley asked whether a Spanish-language version of NIH Medline*Plus* magazine was in the works, and Ms. Kitendaugh replied that it would come out in 2008. Mr. Chabran noted that NLM is on its way to reaching Latinos, with great promotional materials about MedlinePlus, but it should explore expanded usage of them. The CDs about the site are great for those with low literacy, and the tutorials on the M+ esp Web site are also very helpful to that group. Ms. Frierson commented that 40% of U.S. counties still lack broadband Internet access, so print items, toll-free numbers and other "offline" promotional methods still have a place. Dr. Morton thanked the group for a passionate discussion on how to reduce health disparities and help people take greater responsibility for their own health care.

XII. TURNING THE PAGES UPDATE

History of Medicine Director Dr. Elizabeth Fee announced that, with assistance from Guy Cobolet, director of the Bibliothèque Interuniversitaire de Médecine in Paris, "Turning the Pages" is now available in French, as "Tournez Les Pages." Dr. George Thoma of the Lister Hill Center was Monsieur Cobolet's principal collaborator. (Spanish and Portuguese versions are also in the works, and others may follow.) Dr. Fee demonstrated the French language version, and then the newest, and also the oldest, addition to the popular online "Turning the Pages" project, the Smith Papyrus. This work, the world's oldest surviving surgical text, was created about 1600 B.C. and now resides in the collection of the New York Academy of Medicine. It is 15 feet long and includes information on 48 different cases, written in cursive hieroglyphics. The front of the papyrus is devoted to rational medicine, while the reverse presents spells and the more magical side of ancient Egyptian medicine. Dr. Fee credited Michael Chung for coordinating the graphics and animation, Dr. Glenn Pearson for the computer program and Michael North for the historical research. Kathryn Mendenhall of the Library of Congress commended the NLM team and suggested that *today's* treatments for the ailments in the papyrus's 48 cases be added to the site.

XIII. NLM GRANTEE PRESENTATION: THE CIVIL WAR AND AMERICAN MEDICINE

Dr. Margaret Humphreys, Professor in the History of Medicine at Duke University, discussed her research under an NLM publication grant. She highlighted the significance of the Civil War — it ended slavery, devastated the South, killed over 600,000 men, led to the dramatic growth of the federal government and had a powerful impact on the practice of medicine. Among the medical advances that came in the war's wake were: care in hospitals for all; medical specialization; greater prominence for research; the concept of a standard of care; an awareness of health disparities (in this case, between black and white troops); the emergence of nursing as a profession; improved pharmaceuticals; germ theory, which led in turn to aseptic surgery; and the development of public health as a field. Dr. Humphreys has already published one book (*Intensely Human: The Health of the Black Soldier in the American Civil War*) and she expects to produce another in the not-too-distant future. She thanked NLM for its generous support.

XIV. REPORT FROM THE SUBCOMMITTEE ON OUTREACH AND PUBLIC INFORMATION

Ms. Eileen Stanley, Chair of the Subcommittee on Outreach and Public Information, gave a report on the minutes of the Subcommittee's September 18 morning meeting. Ms. Stanley said the Subcommittee heard about the distribution of the NIHMedline*Plus* magazine. It is celebrating its one year anniversary and has proven to be quite successful. The next issue is due out next week. It will feature Kirk Douglas and deals with the issue of stroke which is very timely. It is utilizing Mr. Douglas's experiences because he has just written a book and has experienced a stroke himself. The discussion surrounded issues of increasing distribution. The

goal is to reach a distribution of a million. We started with 14,000 copies and are currently distributing 300,000 copies with a readership of about two and a half million people — so we are on the way! We discussed the idea of affiliations with medical associations and Dr. Ball brought up the idea of increasing outreach activities with nursing groups, particularly because of their role in the education and discharge process of patients – leading from inpatient to outpatient and home care. And we heard that a Spanish version of the magazine is expected out next year. The next item discussed by the Subcommittee was the status of the Go Local Projects. There were two new Go Local projects that took off this summer and we had some discussion about how to tie the project to educator tools to be used in educational settings observing that it is good to not only get information out there but also information about what services are available as well. Lastly, the Subcommittee heard about ENIOP's meeting in South Dakota which brought NLM together with the tribal college. The meeting sought to determine how to respond to the NLM Long Range Plan and to determine how to bring exhibits to native populations. We have companion materials that are already out there but exhibits are also desirable.

XV. NIH POLICY FOR GENOME-WIDE ASSOCIATION STUDIES

Dr. Teri Manolio from the NHGRI, formally of the NHLBI, talked about the NIH policy for genome-wide association studies – known as GWAS. There has been an effort by NCBI for the last year and a half to develop a database that would accept genetic data from GWAS studies. NIH is interested in advancing genome-wide association studies to identify common genetic factors that influence health and disease, and believes that the full value of GWAS can be realized only if the genotype and phenotype datasets are made available as rapidly as possible to scientific investigators. This really builds upon the experience of the Human Genome Project in which sequence data was made available. The Framingham study has also been sharing information for some time, since the late 60s. There have been other clinical data sets distributed over time and other genetic datasets have been available as well. Some are asking why the sudden interest in GWAS? The answer is that the cost of doing them has dropped from about $20 billion in 2001 to about $1 million today. Dr. Manolio discussed the proposed policy, provided the data management overview, submission procedures, access procedures, and protections for research participants, plus issues related to publication.

All investigators who receive NIH support to conduct genome-wide analysis of genetic variation in a study population are expected to submit descriptive information about their studies for inclusion in an open-access portion of dbGaP. She mentioned the Framingham study data are now available with 13,000 variables. The genotyping data will be available for download soon. Everyone is trying to make this as seamless as possible for investigators to access. If someone wants to submit a dataset to dbGaP (a "PGWAS repository"), the funding institution has to certify approval of the submission. When they certify approval they have to include a statement that the data are provided in accordance with the applicable laws and regulations of their state and that an IRB has reviewed the submission. We need information on any limitations on the use of the data, specifically we need informed consent. People who participate in such studies usually provide consent for limited use of the data. NIH is trying to get these consent forms

written more liberally so the data can be used for more purposes. Participants vary widely in their desire or lack of desire to have their personal genetic information used. There are some concerns about the commercial use of the data. The database, dbGaP, doesn't need IRB approval as long as identifying information is stripped by investigators and institutions have agreed that identities of participants will not be disclosed. There are two levels of access to the data once deposited in dbGaP. The basic descriptive information for each NIH-supported GWAS will be available to the public through the open-access division of dbGaP. Access to the individual-level genotype and phenotype data, however, will require authorization from an NIH Data Access Committee (DAC). Investigators and institutions seeking individual-level data from the GWAS data repository will be expected to meet data security measures and asked to submit a Data Use Certification. This certifies that investigators will use the data only for the approved research use, protect data confidentiality, follow appropriate data security protections, follow all applicable laws and any local institutional policies and procedures for handling GWAS data, not attempt to identify individual participants from whom data within a dataset were obtained, not sell or share any of the data to anyone other than those listed in the request, acknowledge the GWAS policy with regard to publication and intellectual property, and provide annual progress reports on research using the GWAS dataset. Committees have been established to review controlled access requests. When a requestor gets approved for access they are given a login password and are authenticated. dbGaP will track the application and approval status and let the applicant know when it is okay to download the datasets. Publication is an issue and is a concern – so there is a period of exclusivity that will apply to those who produced the data.

Dr. Manolio said implementation has been a challenge but all are working incredibly hard to make the repository available. Dr. Morton thanked Dr. Manolio for her leadership in this area. NCBI's Dr. Jim Ostell said it was a bit of a balancing act – but the final dataset gets to the investigator at the same time it gets to everybody else. Dr. Morton also asked if people can withdraw at any time from a study – when data has been released. Dr. Manolio said institutions and submitting investigators can withdraw people before the data are made available. People can even be withdrawn after the data have been submitted to dbGaP – but it is complicated. Once the data are out there for awhile, you can't take them back. Dr. Ostell said that there are cases where people have withdrawn from studies – it tends to be a small number of people. There have been studies where results have been published and then participants have withdrawn their consent – you can't pull back the publication or the analysis or pull back the dataset at that point. We can remove them from the public site and not show the old version, but can't do much beyond that. Dr. Rossiter asked if any of the databases have service levels and costs. Dr. Manolio said that some may have that. Medicaid databases do that. Dr. Rossiter also asked if there is anything that the NLM can do to help. Dr. Manolio said that NCBI is already doing it. Dr. Lindberg asked if any of these studies are clinical trials. Dr. Manolio said not yet, but they are encouraging such studies to be included. Dr. Lindberg said he asked the question because there is a completely different set of legal requirements that are attached to clinical trials and those are changing with respect to registration and reporting results which they are requiring NLM to do. Dr. Morton wondered as studies are merged – how do you deal with phenotypes

that are already fuzzy, like ADHD? Dr. Manolio said that they have to come up with something that can be used and understood across studies like those involving ADHD.

XVI. ADJOURNMENT

The Board of Regents meeting was adjourned at 12:00 p.m. on September 19, 2007.

ACTIONS TAKEN BY THE BOARD OF REGENTS:

- Approval of the May 8-9, 2007 Regents Minutes
- Approval of September 16-17, 2008 Meeting Dates
- Establishment of the Working Group on Clinical Trials

Appendix A - Roster - Board of Regents

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

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

Cynthia C. Morton, Ph.D.
Chair, NLM Board of Regents