DEPARTMENT OF HEALTH AND HUMAN SERVICES
NATIONAL INSTITUTES OF HEALTH
NATIONAL LIBRARY OF MEDICINE
NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION

PUBMED CENTRAL NATIONAL ADVISORY COMMITTEE

Function of the PubMed Central National Advisory Committee

Since the mission of NIH is to conduct and support medical research and to disseminate the results of that research widely to the public and the scientific community, it will make use of electronic publishing technology to fulfill this role by establishing and maintaining PubMed Central. This new service is a Web-based repository, housed at the NCBI that will archive, organize, and distribute peer-reviewed reports from journals in the life sciences, as well as reports that have been screened but not formally peer reviewed. The Committee shall advise the Director, NIH, the Director, NLM, and the Director, NCBI, concerning the content and operation of the PubMed Central repository. Specifically, it is charged to establish criteria to certify groups submitting materials to the system, monitoring the operation of the system, and ensuring that PubMed Central evolves and remains responsive to the needs of researchers, publishers, librarians and the general public.

Summary Minutes of Meeting – October 20, 2005

The meeting of the PubMed Central National Advisory Committee was convened on October 20, 2005 in the Board Room of the National Library of Medicine (NLM), Bethesda, Maryland. The meeting was open to the public from 9:30 a.m. to 2:50 p.m. Mr. James Williams presided as Chair.

Members Present
Shirley Baker, Washington University
Anthony Delamothe, M.D., British Medical Journal
Heather Joseph, M.A., SPARC
Samuel Kaplan, Ph.D., Houston Medical School
Robert Kiley, M.Sc., Wellcome Trust
Debra Lappin, J.D., Princeton Partners Ltd.
Bob Roehr, B.A., Self-Employed
Mary Ryan, MLS, University of Arkansas Medical Sciences
Anthony So, M.D., Duke University
James Williams, M.S., University of Colorado at Boulder
David J. Lipman, M.D., Director, National Center for Biotechnology Information, NLM, NIH, and PubMed Central National Advisory Committee Executive Secretary

NLM Staff Present
Jeff Beck, IEB, NCBI
Dennis Benson, Branch Chief, IRB, NCBI

Jane Davenport, IEB, NCBI
Brooke Dine, IEB, NCBI
Mark Desierto, IEB, NCBI
Jason Eshleman, IEB, NCBI
Martha Fishel, PSD, LO, NLM
Marla Fogelman, IEB, NCBI
Demian Hess, IEB, NCBI
Betsy Humphreys, Deputy Director, NLM
Jennifer Jentsch, IEB, NCBI
Laura Kelly, IEB, NCBI
Donald King, Deputy Director, Research and Education, NLM
Sheldon Kotzin, Chief, BSD, NLM
Sergey Krasnov, IEB, NCBI
David Landsman, Branch Chief, CBB, NCBI
Donald A.B. Lindberg, Director, NLM
Dawn Lipshultz, NCBI
Becky Lyon, LO, NLM
Adeline Manohar, IEB, NCBI
Jim Ostell, Branch Chief, IEB, NCBI
Edwin Sequeira, IEB, NCBI
Elliot Siegel, NLM
Kent Smith, Contractor, NCBI
Jack Snyder, SIS, NLM
Tim Valin, NCBI

Visitors Present
Laura Brockway, Federation of American Societies for Experimental Biology
Norman Frankel, Director, American Medical Association, Allied Publishing Activities
Richard Johnson, Senior Advisor, SPARC
George Kendall, Manager, Proceedings of the National Academy of Sciences
Rebecca Kennison, Director, Public Library of Science
Alice Ra'anan, American Physiological Society
Michael Rogawski, NINDS, NIH
Beth Rosner, Director, American Association for the Advancement of Science, Office of Publishing and Member Services
Crispin Taylor, Executive Director, American Society of Plant Biologists
Nancy Winchester, Director, American Society of Plant Biologists, Publications

I. Call to Order and Opening Remarks
The meeting was called to order at 9:35 a.m. Mr. Williams welcomed members of the PubMed Central National Advisory Committee. Minutes from the April meeting were approved. The next PMC Advisory Committee meeting will take place on Wednesday, April 26, 2006.

II. Upcoming Membership Changes
Six members of the committee will be leaving after this meeting, James Williams, Mike

Eisen, Tony Delamothe, Paula Kaufman, Gerry Rubin, and Ajit Varki. Betsy Humphreys notified Mr. Williams that nominations for six new members have not been given clearance at this time but NCBI will be notified upon approval.

III. Remarks by NLM Director

Dr. Donald Lindberg thanked members of the PMC Advisory Committee for their time and dedication. He also mentioned that parting members will be missed. He noted the NLM budget is still not approved, but he assured the Committee there will be support for PMC. NIH funds overall will be reduced. Referring to the order of business for the day, he agreed with improving standards for PMC and the need for encouragement of new participating journals.

Dr. Lindberg was asked to update the Committee on NLM's Clinical Trials database. He gave a brief background into the inception of the database, starting with NIH trials, then opening submissions to the private sector. Due to some publicized cases involving adverse drug effects, a statement was issued by the International Committee of Medical Journal Editors that study results would only be published from trial reports that are well described and publicly available. This decision has resulted in large depositions into ClinicalTrials.gov, about 6,000 in the last month alone. Mr. Williams asked if international trials are accepted into the database. Dr. Lindberg responded that NLM will register international trials and an international health authority, such as WHO, will be identified to provide registry numbers for reporting, similar to FDA registration numbers for U.S. trials.

An update on the PubChem chemical database was also requested. Dr. Lipman reported that it has over 9,000 users daily, currently contains over 4 million compounds and initial bioassays have been submitted by the first NIH-funded screening center. An NIH request was published in the Federal Register inviting private sector providers and users of chemical information to serve on a working group of the NCBI Board of Scientific Counselors that will advise on interactions with private sector chemical information providers in the development of PubChem. The initial meeting is planned in December.

Ms. Lappin asked for a perspective on a higher level vision for new tools and potential effects on research. Mr. Roehr asked about the possibility of Clinical Trials being a repository for raw data. Dr. Lipman responded that some experimental gene information is publicly available, and NHLBI has a policy for new clinical studies adding anonymized data to a publicly available database. NIH-wide, there is an effort for more transparency and access to clinical information as well as integration of the underlying data and related research information. Dr. Lipman noted that PubChem and ClinicalTrials.gov will have strong implications for understanding the molecular basis of disease and drug development.

Mr. Williams asked if there is a parallel between submission processes in PMC and ClinicalTrials.gov as an incentive-based model to get more data into the databases. Dr. Lindberg mentioned that the motivation must come from the investigator. The Committee was interested in staying informed on developments in PubChem and Clinical Trials

since they have impact on access to the published journal literature.

Dr. Kaplan suggested the possibility of job-fair type of meeting for publishers to recruit more PMC participants. Dr. Lipman emphasized that it must be made clear that the open house would be for publishers who will participate directly in PMC, rather than for submissions complying with the NIH Public Access policy where the author must be directly involved. Dr. Lindberg added that important aspects of PMC such as preservation of data and back issue digitization need to be relayed to publishers. All Committee members agreed that a "fair" for publishers would be a good idea. Dr. Lipman and Dr. Kaplan will discuss the topic further and report at the next meeting.

IV. Revised Acceptance Policy for New Journals

At the April meeting there was a discussion about using NLM's Literature Selection and Technical Review Committee (LSTRC) to review the quality of any non-Medline journals that apply for PMC participation. At that meeting, the PMC committee asked for more details on LSTRC's role in a new review process and agreed to using LSTRC as needed to review new journals in the interim.

Mr. Ed Sequeira explained that the original PMC acceptance criteria stated a journal must be either in Medline or one of the other major abstracting and indexing databases, or it must provide letters of support from three members of its editorial board who currently are principal investigators on research grants from major funding agencies such as NIH. The new process, introduced on a trial basis subject to committee approval, adds two conditions for non-Medline journals. If the journal publisher already has other journals in Medline or PMC, the new journal will be approved for PMC on this basis, subject to a possible later review by LSTRC should questions arise about journal quality. Otherwise, the journal is asked to provide details of its editorial board, peer review process, and scope of information. LSTRC will review this information and the quality of recent articles and make a recommendation.

Mr. Sheldon Kotzin provided background information on LSTRC's function for Medline. For PMC reviews, a decision on a journal is expected within two weeks. He added that the standards for PMC journals will be less exclusive than those for Medline. Dr. Lipman explained that the LSTRC review approach will provide a balance between participation, quality control, and efficiency.

Dr. So reported on a recent trip to Cape Town, South Africa, where he learned that African publishers often feel slighted by western publishers and information databases. He inquired about their representation in PubMed and the possibility of outreach. Mr. Kotzin replied that Medline is actively seeking journals from sub-Saharan Africa that focus on local issues. In fact, four African journals are currently being mentored by four western journals for infrastructure and quality development for indexing in Medline.

Dr. Kaplan asked if a process exists for a retrospective look at a journal after a certain period of time. Mr. Sequeira answered that a method is not set up at this time, but could be in the future. Mr. Kiley asked if there is concern that the public access submissions

could change the quality of PMC. Dr. Lipman said that there have been no problems so far because most articles submitted are published in high quality journals.

A motion was carried by the Committee to accept the revised PMC acceptance and review policy.

Break 11:07-11:15

V. Update on Back Issue Digitization Project

Ms. Carol Meyers who manages the back issue digitization project for PMC began her presentation with a brief background on the project. The project started in September 2002, and now has roughly 100 titles in various stages of production. Currently, the *Bulletin of the Medical Library Association* back to 1911, all ASM journals and predecessor titles back to 1916, as well as *The EMBO Journal*, *Nucleic Acids Research* to 1974, and the *Proceedings of the National Academy of Sciences* to 1915 have been digitized and are publicly available. British Medical Association specialist titles are completely scanned and will be live via a phased implementation program over the next year.

Journals still in production include the *American Journal of Public Health*, which has been accumulated back to 1873, the *Canadian Veterinary Medical Association*, and *Public Health Reports*. Ms. Meyers discussed the Wellcome Trust collaboration under which the *Medical History* journal archive was completed in February 2005 and the *Biochemical Journal* will be completed by the end of the year. Thirteen additional titles are in the Wellcome Trust pipeline.

A chart of progress to date illustrated the large amount of effort involved in the project, such as tracking journals and obtaining archival issues. Challenges include quality assurance and information tracking. A style guide has been created for each journal to assist the scanning process. It was asked if the process is physically destructive to the paper copy of the journal. It is, but NLM retains its own archival copies of all journals. Donors provide the copies that are scanned. Maximum scanning production by the contractor is about 300,000 pages per month.

The file of some completed archives have been delivered to their publishers, such as the ASM journals and Biochemical Journal. Mr. Williams asked if the publishers are creating their own archives. Ms. Meyers responded the ASM and Biochemical Journal archives will go on their own sites. Ms. Johnson and Mr. Kiley noted that the project provides a tremendous service of great quality with added value such as uniform citation creation and reference extraction. Mr. Roehr asked about the capacity of the project. Dr. Lipman answered that if many publishers wanted to sign on, it would not be a problem. A discussion about obtaining issues from donors ensued. Dr. Delamothe suggested the need to get the project and its value communicated to the publisher community. Dr. Kaplan agreed that ASM will do an article on the benefits of the project and encourage other experienced publishers to do the same. Mr. Williams commended the projects and added that NCBI and NLM are advancing the future of digital scholarship.

VI. NIH Public Access
Dr. David Lipman reported on a July 11 meeting of the NIH Public Access Working Group meeting formed by the NLM Board of Regents. The working group which advises on the implementation of the NIH Public Access Policy has representation from a broad range of interested stakeholders including researchers, publishers, societies, libraries, and the public. Minutes from the meeting are available on the NLM web site and the next meeting will be held on November 15.

Dr. Lipman reported on submission rates to the NIH Manuscript Submission System (NIHMS). Each month only about 3.8% of eligible NIH publications are being submitted. Dr. Lipman noted that 60% of submissions so far have opted for no embargo, and only 19% have chosen the maximum of 12 months.

Dr. Lipman explained that the actual submission process takes only about 10 to 15 minutes on average. Mr. Williams asked about bulk submissions from publishers. Dr. Lipman replied that third party submissions are currently limited to single submissions. A bulk submission system will be available for publishers in the future, but authors and PIs will still be required to verify the submission. There was discussion on ways to encourage greater participation.

Lunch 12:30-1:00

VII. Update on Wellcome Trust Open Access Policy
Mr. Robert Kiley discussed the Wellcome Trust policy which states that as of October 1, 2005, all new grant holders will be required to deposit research papers into PMC and existing grantees will be urged to deposit. As of October 1, 2006, the policy will apply to all grant holders, regardless of grant date. Mr. Kiley thanked David Lipman and PMC for tailoring the NIH manuscript submission system for Wellcome grantees. Wellcome will meet submission costs for deposition into open access journals, which will total about 1-2% of Wellcome's research budget ($350 million per year). He reported that some publishers have modified their policies to comply with the new policy.

Three primary reasons for the Wellcome Trust open access policy include: long term preservation of information, links to underlying data, and links to the grant system for information analysis and feedback on funding. Efforts are being made to speak to the research community and highlight benefits of open access.

Dr. Lipman noted that working with the Wellcome Trust on this project has been a good learning experience and provides a model for other funding organizations. A question was asked regarding conflict between the publisher copyright agreement and the Wellcome policy. Mr. Kiley answered that the grant agreement would predate the publisher agreement, by years in some cases. He explained that some suggested solutions to avoid conflict are: discussing policies with the journal editors in advance of submitting a paper to a journal, using a license to secure the access permissions, supplying additional text regarding intentions along with the agreement, and lastly, finding a different

publisher. Dr. So asked if Wellcome could assist other groups with setting up similar programs. Dr. Kiley replied that at the moment, their focus is to work with UK-funded groups. A discussion ensued regarding promotion of open access policy and verification and correction of data. The Committee applauded Wellcome's open access policy and collaboration with PMC.

VIII. Evolution of the PMC and NIH Manuscript Submission System (NIHMS) Infrastructure

Dr. Ostell began by giving a brief definition of NCBI's bibliographic resources ranging from portable PMC and the NLM DTD to NIHMS. PMC first developed a way for conversion of content in various formats to standard XML format, with storage and rendering in one format. The Harvard e-journal archiving project led to a collaboration that expanded the PMC DTD to accommodate journals from disciplines other than the life sciences. The resulting two NLM journal article DTDs include, (1) an Archiving and Interchange DTD designed to preserve intellectual content and written for ease of conversion and completeness, and (2) a Journal Publishing DTD which is a fully compatible subset of the Archiving DTD. Groups such as Highwire Press, JSTOR, Public Library of Science, and other PMC contributors have adopted the NLM DTD. Dr. Lipman added that these tools demonstrate the research and development aspect of NCBI that result in products for others to use as they wish.

Portable PMC (pPMC) was designed so that a mirror of PMC could be made available in other countries. Test versions of pPMC have been running on three sites in the UK, Italy, and South Africa. Dr. Ostell provided an explanation of the pPMC modules and how data searching, collection, and rendering are performed.

The NIHMS process for authors to submit manuscripts in compliance with the NIH Public Access policy was demonstrated. Tools, such as a portable renderer for converting XML to HTML, have been developed which are available for adoption by other publishing groups.

The NCBI Bookshelf was initially developed to tag and digitally convert scientific texts offered by publishers for access on the NCBI web site. Over the years, newly published collections have been added and now books are being created for the Bookshelf. One book has been released first on the Bookshelf, ahead of print. Some books have been authored by NCBI using an XML authoring system built upon MS Word. An experiment is now underway to incorporate books into the NLM DTD.

In early 2006 a new version of pPMC will allow searching locally and reduce bandwidth for updates, important for countries with limited internet connectivity. About mid-2006, a portable NIHMS is expected to be completed that will allow collaborators to locally process the submissions. In late 2006, the entire process will be performed locally with addition of other NLM DTD content to pPMC.

Dr. Ostell explained that the ultimate vision, a "harmonic convergence" of resources and software, is to get information from publishers in the NLM DTD so that no conversion is

needed. Dr. Lipman added that the idea is to have an infrastructure compatible with formats that will allow for inexpensive and flexible data exchange as well as local archive building.

Dr. Kaplan asked if books could be added to the Bookshelf since ASM has monograph collections. Dr. Lipman replied that NCBI would welcome the ASM monographs. Dr. So asked if sales of Bookshelf books have gone up or down. So far, there is no indication that book sales have been hurt. Ms. Humphreys added that one group said that previously published titles have had an increase in sales. It was also asked if video clips can be shown. There are video clips available currently in PMC and the books can be interactive. Mr. Kiley asked if there are plans to digitize early books. The History of Medicine Division of NLM is doing some early book digitizing which will be available in the Bookshelf. However, NCBI is primarily interested in making online versions of books based on an XML encoding of the content, rather than simple digital scans.

Ms. Lappin asked about NLM's potential role in analyzing research portfolios, based on comments from Dr. Zerhouni regarding raw data and PubChem. Ms. Humphreys reported that some NIH projects for portfolio management are using NLM resources. Dr. So asked if any institution will eventually be able to get pPMC noting that the vision of sharing data and creating archives is important. Dr. Lipman replied that the model software system would be usable with other data over time for outside repositories, but may not be titled PMC when other content is used.

Dr. Lipman thanked the retiring members who were present, Tony Delamothe and Jim Williams for their service to the Committee.

IX. Adjournment

The PubMed Central National Advisory Committee adjourned the public meeting at 2:50

CERTIFICATION

I hereby certify that the foregoing minutes are accurate and complete.

_____	_____
James Williams, Chair (Date)	David J. Lipman, M.D., Director (Date)
PubMed Central National Advisory Committee	Director, National Center for Biotechnology Information, NLM