

DEPARTMENT OF HEALTH AND HUMAN SERVICES

NATIONAL INSTITUTES OF HEALTH

PUBMED CENTRAL NATIONAL ADVISORY COMMITTEE

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION

NATIONAL LIBRARY OF MEDICINE

March 21, 2001

**NLM Board Room
National Center for Biotechnology Information
National Library of Medicine
8600 Rockville Pike
Bethesda, Maryland 20894**

**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 –MARCH 21, 2001

The meeting of the PubMed Central National Advisory Committee was convened on March 21, 2001 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 4:00 p.m. Dr. Joshua Lederberg presided as Chair.

Members Present

Joshua Lederberg, Ph.D., The Rockefeller University, PubMed Central National
Advisory Committee Chairman
Patrick O. Brown, M.D., Ph.D., Stanford University
Nicholas Cozzarelli, Ph.D., University of California at Berkeley
Uta Francke, M.D., Stanford University Medical Center
Frank Davidoff, M.D., *Annals of Internal Medicine*
Paul Ginsparg, Ph.D., Los Alamos National Labs
Michael Homan, Mayo Medical Center Libraries
Elizabeth Marincola, American Society of Cell Biology
Suzanne McInerney, Health Writer/Patient Advocate
James G. Neal, Johns Hopkins University
Maurice F. Rabb, M.D., University of Illinois at Chicago
Vitek Tracz, Current Science Group
Harold Varmus, M.D., Memorial Sloan Kettering Cancer Center
James F. Williams, University of Colorado at Boulder
David J. Lipman, M.D., Director, National Center for Biotechnology Information, NLM, NIH,
PubMed Central National Advisory Committee Executive Secretary

Invited Guests Present

Tony Delamothe, Web Editor, BMJ Publishing Group

John Inglis, Director, Cold Spring Harbor Laboratory Press

Nancy Winchester, Publications Director, American Society of Plant Physiologists

Melissa Junior, Managing Editor, *Plant Physiology*, American Society of Plant Physiologists

Beth Staehle, Managing Editor, *The Plant Cell*, American Society of Plant Physiologists

NLM Staff Present

Donald A.B. Lindberg, Director, NLM

Kent Smith, Deputy Director, NLM

Steven J. Philips, Assistant Director for Research and Education, NLM

Betsy Humphreys, Associate Director for Library Operations, NLM

I. Call to Order and Opening Remarks

Dr. Lederberg welcomed members of the PubMed Central National Advisory Committee. The Committee officially adopted the minutes from the previous meeting and agreed that the fall meeting would be Wednesday, October 10, 2001. A tentative date of March 20, 2002 was set for the spring meeting.

II. Remarks by NLM Director

Dr. Lindberg thanked Dr. Lederberg for the opportunity to address the Committee. He next requested that the Members consider potential solutions to the problem of how to preserve and archive the numerous visual data formats in existence today and to consider working in partnership with other organizations involved in similar projects. The Committee agreed to add this item to the October meeting agenda. Deliberation ensued, including an informal discussion of NCBI's role in, and current strategy for, converting data/information to a single, common XML DTD format. Dr. Lipman iterated that PubMed Central (PMC) is a work-in-progress. Current efforts focus on creating a common text format that combines the concepts of utility, stability, and sustainability, but NCBI would consider options for the long-term preservation of visuals. **Action Item—Members should be prepared to discuss further the issue of long-term preservation of data at the October 2001 meeting.**

Dr. Lindberg invited Committee members to view NLM's exhibit "Turning the Pages," a program developed at the British Library. "Turning the Pages" uses computer animation, high-quality digitized images, and touch screen technology to simulate the action of turning the pages of the book *Elizabeth Blackwell's Curious Herbal*. Dr. Lindberg cited this exhibit as an example of a long-term "digital collaboration" beginning with the British Library's adoption of NLM's MEDLINE technology.

Discussion shifted to NLM's role in digitizing print format archives. Dr. Varmus reminded the Committee that a Back Issue Scanning Working Group had been formed at the previous meeting and consisted of Harold Varmus, Patrick Brown, and Vitek Tracz. **Action Item—Members of the Working Group agreed to convene and discuss issues such as the appropriate number of back-years to include in the PMC archive, which journals to include, and the number of journal issues. The Working Group should be prepared to summarize their discussion at the next meeting.**

The Advisory Committee felt it would be helpful to hear perspectives from other organizations involved in similar digital conversion projects. **Action Item—Dr. Lipman agreed to invite Kevin Guthrie, Executive Director, JSTOR to speak at October’s meeting.**

III. Need for Enhanced Outreach and Education Efforts

Dr. Lederberg called to the Committee’s attention a recent correspondence outlining several perceived problems with PMC progress, including the inability to reliably “search” PMC. Dr. Lipman responded directly, indicating that members of the PMC development team recently enhanced the capacity to search the archive by introducing a more meaningful search “vocabulary”. Using this expanded vocabulary, users may search either an entire journal article—not just the abstract or MeSH indexing terms, or the entire archive, regardless of the journal.

Concern was expressed about the pace at which PMC is progressing. In fact, NCBI has taken a deliberate approach to ensure a stable and consistent format for storing and retrieving text. This process of conversion to a common PMC DTD is critical to achieving a viable, long-term archive—a key goal of PMC. NCBI has made major efforts towards improving PMC operations, primarily by introducing changes that streamline the process by which PMC validates publisher-submitted SGML/XML. In the long run, these changes will provide a stable foundation for building and maintaining the PMC archive.

Lastly, it appears there may be public confusion surrounding the terms “central repository” versus “single repository.” Some believe NCBI is promoting PMC as the sole source for archiving peer-reviewed literature, when, in actuality, NCBI endorses wide dissemination of PMC via alternate or mirror sites. It is the concept of disseminating information from multiple sites that ensures the utility, stability, and sustainability of PMC. Dr. Lipman did mention that by accessing PMC via NCBI’s site, a user is guaranteed a level of uniform functionality and can take advantage of NCBI’s integration of PMC with other biological information resources.

The Committee recognized that these perceived problems, as well as the threat of potential loss of revenue, were preventing many publisher organizations from joining PMC. To encourage wider publisher participation, an option was proposed that would require full submission of published material to PMC, but would allow the viewing of full-text at the publisher’s site. Specifically, the publisher would agree to furnish the NLM with full-text articles in either SGML or XML format that conforms to an established DTD. NCBI would transform this data into a PMC DTD and load the data to create the PMC archive. From this common archival format, standard, as well as specialized, indices would be built to facilitate end-user searching of journal articles. Articles retrieved in searches however, would have links to the publisher’s site where the user would then be able to view the full-text.

Under this option, the publisher would agree that material older than a determined interval—preferably six months, but no greater than 12 months—would be viewed at no cost to the user. Material newer than the determined interval could be available from the publisher’s site for a fee. This option would carry a restriction that, if at any time, the publisher removed free access to the older material, PMC would be permitted to provide free full-text to this material at the PMC site.

It was agreed that PMC would continue to actively promote free access of material at the PMC site as the preferred mode of publisher participation for the obvious benefits of searching and viewing at a single site. The optional level of participation was viewed as a pragmatic attempt to

broaden the number of participants, enrich the body of searchable text, and, most importantly, create a functional archive. **Action Item—The current “PubMed Central Participation Agreement” will be re-drafted to incorporate the terms of the new participation option.**

The Committee convened for lunch from 12:00-1:00

IV. PubMed Central Update

Dr. Lipman provided the Committee with an update on PMC’s status. He demonstrated for the Committee the enhanced capacity to search and retrieve relevant information from within the PMC archive and the ability to link to additional NCBI resources, such as OMIM and the three-dimensional protein structure database. Dr. Lipman also discussed several recent improvements to the PMC site designed to increase the ease of user navigating. Other topics covered in Dr. Lipman’s presentation included:

- Journals available online;
- Journals that have agreed to participate in PMC;
- Journals currently in the data evaluation stage;
- Recent revisions and additions to the PMC Web site;
- Utility of a common PMC DTD format; and
- Relevant Web site usage statistics.

Lastly, Dr. Lipman introduced various members of the PMC development and operations team and thanked them for their role in building, supporting, and enhancing the PMC system.

Action Item—In response to the Committee’s view for increased outreach and education efforts, and based on the outstanding progress recorded in Dr. Lipman’s presentation, Dr. Varmus recommended that NCBI prepare a short essay outlining current PMC status. The report would serve to educate the community about PMC, accentuate recent technological advances, and inform the community regarding recent policy changes. The goal would be to draw the community, and more importantly potential contributors, to the site as well as demonstrate the non-competitive nature of PMC.

VI. Conclusion

Dr. Lederberg thanked Committee members for their participation. He restated that the next meetings are scheduled for Wednesday, October 12, 2001 and Wednesday, March 20, 2002, the latter pending agreement by Committee members. Drs. Lederberg and Lipman thanked both the Committee members and invited guests for their valuable time and input.

Committee members noted the benefits associated with inviting guest publishers and suggested that Dr. Lipman continue this effort. Suggestions were also made to have various publishers present at upcoming meetings; that NCBI actively pursue opportunities to present on PMC at outside meetings; and that NCBI communicate with external organizations involved in similar projects to improve public relations.

Ms. Marincola stressed the need for increased outreach and education efforts and once again encouraged the NCBI, as well as the PMC Committee members, to take an active role in promoting the current utility and future potential of PMC.

VI. Adjournment

The PubMed Central National Advisory Committee adjourned the public meeting at 3:30 p.m.

CERTIFICATION

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

(date)
Joshua Lederberg, Ph.D., Chair
PubMed Central National Advisory Committee

(date)
David J. Lipman, M.D., Director,
National Center for Biotechnology
Information, NLM

CONCURRENCE

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