

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

**MINUTES OF THE PUBLIC ACCESS WORKING GROUP
OF THE BOARD OF REGENTS
July 11, 2005**

The first meeting of the Public Access Working Group of the National Library of Medicine's Board of Regents was convened on July 11 at 10:00 a.m. in the NLM Board Room, Building 38, National Library of Medicine (NLM), National Institutes of Health (NIH), Bethesda, Maryland. The meeting was open to the public.

MEMBERS PRESENT:

Dr. Thomas Detre, Dr. Deanna Marcum, Dr. Jeffrey M. Drazen, Dr. Mark Scott Kamlet, Dr. T. J. Koerner, Brian Nairn, Dr. Mark E. Sobel, Michael Stern, Sharon F. Terry, Patricia Thibodeau, Dr. Annette Thomas, Donald E. Tykeson, Dr. Gary E. Ward, Ann Wolpert

EX OFFICIO MEMBER PRESENT:

James F. Williams, Chair, PubMed Central National Advisory Committee

MEMBERS OF THE PUBLIC PRESENT:

Barbara Meredith, Association of American Publishers; Martin Frank, American Physiological Society; Shannon Cummins, American Society of Hematology; Craig Van Dyck, Wiley; Colette Bean, Wiley; Eric Massam, Reed Elsevier, Heather Joseph, SPARC; Prudence Adler, ARL

FEDERAL EMPLOYEES PRESENT:

Dr. Elias Zerhouni, Director, NIH
Dr. Norka Ruiz Bravo, Deputy Director, Extramural Research, NIH
Dr. Timothy Hays, Office of Extramural Research, NIH
Dr. Donald A. B. Lindberg, Director, NLM
Betsy L. Humphreys, Deputy Director, NLM
Dr. Donald W. King, Deputy Director for Research and Education, NLM
Dr. David Lipman, Director, National Center for Biotechnology Information, NLM
Dr. James Ostell, Chief, Information Engineering Branch, NCBI, NLM
Dr. Dennis Benson, Chief, Information Resources Branch, NCBI, NLM
Dr. Bart Trawick, NCBI, NLM
Ed Sequeira, NCBI, NLM
Dr. Sergey Krasnov, NCBI, NLM
Dr. Eugene Yaschenko, NCBI, NLM
Becky Lyon, Deputy Associate Director for Library Operations, NLM
Sheldon Kotzin, Executive Editor, MEDLINE, NLM
Kent Smith, Consultant, NLM
Jane Bortnick Griffith, Consultant, NLM
Suzanne Aubuchon, Office of the Director, NLM
Patricia Carson, Office of the Director, NLM

I. OPENING REMARKS

Dr. Donald A.B. Lindberg, Director of the National Library of Medicine, welcomed the participants to the first meeting of the Public Access Working Group of the Board of Regents. Dr. Thomas Detre, Chair of the Working Group and of the NLM Board of Regents, also welcomed the members and invited them all to introduce themselves. Dr. Lindberg and Dr. Detre provided an explanation that conflict of interest rules governing the Working Group members relate only to the specific Charge of the Working Group. The Charge states that the Working Group will:

1. review statistical evidence on the impact of the policy, e.g., number of manuscripts submitted, summary data on embargo periods, connections to other NIH information resources, level of use, etc.;
2. provide suggestions for improving the implementation of the manuscript submission system and procedures;
3. assess the extent to which the policy is achieving its stated goals; and
4. suggest any changes to the policy that might further these goals.

II. NIH Public Access Policy: Origins, Goals, Provisions

Dr. Norka Ruiz Bravo described the origins of the NIH Public Access Policy, highlighting language that appeared in the House Appropriations Committee Reports for FY2004 and FY2005 supporting NIH's development of a policy to improve public access to the peer-reviewed, published literature resulting from NIH-funded research. Dr. Ruiz Bravo identified the three primary goals of the NIH policy: 1) to create a central archive of NIH-funded research publications, 2) to advance science and enable NIH to better manage its research portfolio, and 3) to provide electronic access to the public to NIH-funded research publications. She also described the main components of the NIH Policy, its coverage, and the timing for submission of manuscripts. Dr. Ruiz Bravo stated that NIH does not consider itself a publisher in the traditional sense, and emphasized that NIH strongly supports the current publishing process by providing direct costs to investigators to cover publication costs and through indirect costs that support journal subscriptions. The NIH Policy is compatible with any publishing model. She reaffirmed that the NIH Policy explicitly recognizes and upholds copyright and that NIH is encouraging investigators to sign agreements that specifically allow their manuscripts to be deposited in PubMed Central (PMC) for public posting as soon as possible after journal publication. Advantages of depositing in PMC include fulfilling part of the NIH grant progress reporting, enabling cross linking to many related databases and resources, heightening visibility and likelihood of citation, and ensuring the permanence of the record of science. Dr. Ruiz Bravo closed by providing the recent language that appeared in the report of the House Appropriations on FY2006 funding for the Departments of Labor, Health and Human Services, and Education. In summary, the Committee directs the NIH Director to submit a report by March 1, 2006 on the progress achieved during the first eight months of the Policy and also directs that an aggressive education and outreach initiative be developed to inform grant recipients about the policy.

Ms. Ann Wolpert asked what is meant by the Policy's request to deposit supplementary data. Dr. David Lipman responded that supplementary material refers to whatever additional material is submitted to a journal in addition to the manuscript. Dr. Jeffrey Drazen asked whether data existed to show that deposit in a public repository actually increases the amount of times an

article is accessed. Dr. Ruiz Bravo stated that there is data that demonstrates this and Dr. Lipman said NLM could provide it. Dr. Annette Thomas questioned NIH's statement that it is not a publisher given the dynamic nature of what might be considered a publication in the Internet environment. Dr. Drazen asked whether the NIH Policy would apply to a paper where only a small percentage of the funding supporting the associated work came from NIH. Dr. Ruiz Bravo replied that it would.

III. STATISTICAL SUMMARY OF SUBMISSIONS SINCE MAY 2, 2005

Dr. David Lipman presented statistics on the number of submissions to date. He began by clarifying that the NIH Policy is requesting the submission of the author's final manuscript *upon acceptance for publication*. Dr. Lipman described the way that data flows through the NIH Manuscript System (NIHMS) and noted that it is similar to the GenBank submission process that grantees have been doing for 20 years. The submission of gene sequences also is voluntary (although many journals require it for publication); however, it requires more work than the manuscript submission system. Dr. Lipman provided data on the interval between publication of a manuscript and its submission to NIHMS. Forty percent of the items deposited to date have not yet been published; 28 percent were published in the last six months, 11 percent were published between 7 and 12 months, and 21 percent were published more than 12 months ago. Using existing data from Medline, it is possible to estimate that on average 5500 manuscripts would need to be deposited monthly (or 250 each workday) to achieve 100 percent participation by NIH funded investigators. Using this as a baseline, the current daily submissions are about 3 percent of unpublished manuscripts, coming from approximately 4 percent of NIH investigators. Other data provided included the fact that 48 percent of the submissions were completed in under three minutes and 68 percent requested that their unpublished manuscripts be made available immediately upon publication. Dr. Thomas asked whether investigators were submitting older, published articles because they were confused about the parameters of the NIH Policy. Dr. Lipman responded that he believed that most of the initial submitters were supporters of the NIH Policy or those who felt obligated to comply with it. These individuals are likely submitting all of their NIH funded articles, including those already published, and indicating immediate public release.

Mr. Tykeson asked how many NIH funded investigators are not submitting manuscripts. Dr. Lipman replied that about 96 percent are not. Dr. Martin Frank asked whether NIH knows the correlation between the delay period indicated by submitters compared to the relevant publisher's policy. Dr. Lipman replied that it is not known, but that the most common place for submitters to stop during the submission process is when they are asked to approve a submission statement acknowledging that they are informing the publisher of their intent to submit the manuscript

Dr. Thomas asked how many of the submissions would already be in PMC. Dr. Lipman stated that it would have been less than 10 percent. Dr. Thomas asked whether NIH would consider listing publishers' policies concerning public release of articles published in their journals to assist submitters. Dr. Lipman responded that the NIH Policy is based on the concept that the authors are responsible for making the determination about the delay period for public release. In addition, publishers' policies are subject to change without notice and are best obtained directly from the source.

Dr. Sobel asked whether PMC journals should have been subtracted from the calculations on the number of articles expected. Dr. Lipman replied that this was done to be conservative in estimating the denominator of anticipated articles. Dr. Ward also noted an author may choose to

make a manuscript available earlier than it otherwise would be for a journal that is already deposited in PMC, and they should be encouraged to do so.

Dr. Detre reminded the Working Group members that NIH was following Congress' direction in developing the Policy.

Dr. Lipman demonstrated how an article that has been submitted under the NIH Policy appears in PMC. He highlighted that it is clearly marked as an author's manuscript and that there are clear links to the publisher's site where the final article can be found. He showed how the tagging of the manuscripts enables links to other resources and more structured searching. The system provides both a PDF version and a Web view that includes all the links. Dr. Frank asked why NIH didn't consider itself a publisher if it was providing these various views of the manuscript. Dr. Lipman responded that NIH is building an active archive and is not operating as a primary publisher. Ms. Wolpert asked whether links are checked to ensure that they are still active. Dr. Lipman replied that one of the values of an archive that is constantly used is that such problems are quickly identified. Dr. Thomas asked about the types of links provided, and Dr. Lipman stated that there are links to many NIH supported databases, and that the number of links is continually growing. He did a demonstration of the system showing a variety of links to different databases.

IV. Stakeholder Comments

Dr. Marcum stated that there is a need to find new models for preserving permanent access to important information resources and that it is important not to lose research materials that may be valuable in the future. She commented that during this transition to more digital information, the concept that libraries are in the best position to be and maintain these archives should be supported.

Ms. Terry commented that she was very disappointed about the small percentage of manuscripts submitted. She believes that the voluntary nature of the NIH Policy puts researchers in a difficult position, which could be eliminated if the policy were mandatory. She viewed what NIH is providing as enhancing science through making connections that otherwise could not be made. She considered this more a part of the scientific enterprise, than acting as a publisher.

Mr. Tykeson asked whether the Working Group needed to wait until after the March report to Congress to recommend changes to the Policy. Dr. Detre replied that it would be helpful if the Working Group could come to consensus on necessary changes and that if that did not happen it was likely that Congress would provide more direction.

Mr. Tykeson stated that he saw a distinction between existing grants that were already awarded and new grants that would be awarded in the future. He recommended that manuscript submission in PMC should be a requirement for any new grants awarded, while for existing grants NIH should make it a request.

Dr. Kamlet stated that the Working Group could be helpful to NIH by determining why the response so far has been so low. Dr. Lipman noted that communication between NIH and its investigators, as well as between investigators and publishers, is often weak.

Dr. Sobel stated that if NIH's goal is to populate the archive, the data for the first few months are not meaningful. He believes that authors are afraid to act until they know what publisher

policies are, and many publishers have not yet had time to develop their journal policies for deposit of manuscripts funded by NIH in PMC and have not yet had time to modify existing copyright transfer agreements.

Dr. Sobel commented that the issue of the definition of what constitutes the final manuscript upon acceptance for publication is a stumbling block and that the existence of multiple versions is dangerous. He recommended that NIH should ask the journals to provide the final version, rather than making the request of the investigators, in order to achieve better quality control.

Ms. Wolpert raised the issue of the role of research and educational institutions in the process. She noted that many academic institutions are trying to put new workflow and other processes in place to assist authors in complying with the policy, but that due to the academic calendar and the time required to accomplish this, it was likely that an impact would not be visible until spring, 2006.

Dr. Drazen stated that he believes the Policy places authors in an untenable position because they do not want to offend either the journals in which they would like to publish or NIH as the sponsor of their research. He is also very concerned about errors in manuscripts and believes that if the mission is to get an accurate archive, then it would be better to work with the publishers, rather than the authors.

Dr. Marcum asked whether any publishers had said they would not participate at all in NIH Policy. Dr. Lipman replied that NLM was unaware of any publishers who had refused outright to participate, but that different publishers have different definitions of what it means to participate. He stated that if a publisher is willing to provide the final manuscript PMC will gladly take it, but the policy was addressed to the authors' manuscripts as accepted after peer review because the authors have original rights in these manuscripts.

Dr. Ward said that while receiving the publisher's final version may be preferable, in the absence of such participation, getting the author's manuscript is the next best thing. He believed that the fact that the Policy was voluntary was the primary reason for non-compliance and that unless authors are strong advocates of open access, they are likely not to submit or to drop out of the submission process when asked to sign the deposit agreement.

Mr. Nairn voiced his concern about rushing to conclusions about whether the policy is working before publishers have an opportunity to submit articles. (The third party submission option was opened on July 6, less than one week before the meeting.) He believes that it will take a year before it is possible to see if the goals are being achieved.

Dr. Sobel commented that there is a significant variation among the types of journals that publish the results of NIH-supported research and the percentage of NIH funded articles they publish. Some non-profit scientific society journals may have 40-70 percent of their content coming from NIH-funded research.

Dr. Drazen stated that *New England Journal of Medicine* was willing to give material to PMC for archival and search purposes, but that journals provide a context that is missing from an archive. Dr. Lipman responded that some publishers have sought a distributed solution where the actual article is available only on the publisher's website. This approach would not enable NIH to create a viable archive. Others have suggested a "dark" archive, where the material would be stored, but not accessible. In order for the archive to be viable for long term use, the

material must comply with PMC's standards and must be regularly used. Without routine use of the contents of an archive, it is not possible to guarantee its usability in the future. This is particularly true for biomedical resources that involve not only text, but extensive data sets.

Dr. Koerner agreed that publishers make an important contribution to the quality and utility of journal articles. He stated that there has not been enough time to evaluate the results of the policy and that the Working Group should set realistic milestones for observing change.

Dr. Drazen added that publishers provide not only context, but also branding that is important to the user of the information. He recommended that while searching the material on PMC was useful, the content should be viewed on the publisher's website in order to get both the context and the impact of the brand. Dr. Lipman responded that unless that information can be used on the PMC site, it will be impossible to maintain quality. He also noted that a link is already provided to the publisher's site for each item in PMC.

Dr. Sobel commented that publishers perform a unique peer review process that is very rigorous and that many publishers have their own online archives and support some form of open access. He suggested that it would be more difficult to get compliance from publishers who don't have any form of open access and that NIH should begin by working with publishers who do.

V. DEMONSTRATION OF NIHMS

Dr. James Ostell gave a presentation of the submission system, showing the steps an author goes through to place a manuscript in NIHMS. He noted how a third party can submit on behalf of an author, but the author must provide final approval and designate the delay period for public access.

VI. REMARKS FROM THE DIRECTOR OF NIH

Dr. Elias Zerhouni thanked the Working Group members for their participation and for providing NIH with guidance and advice in this important area. He noted that there are many valid points of view on this topic, but that he is committed to ensuring that NIH takes advantage of new ways to exchange information and use it in advancing research. Dr. Zerhouni saw the archive being developed as a dynamic tool that is essential to an agency like NIH, with its \$28 billion investment in research. He commented that while many people advised him not to address the issue of scientific publishing, he believes that the status quo must be challenged in a way that makes science available to the public without destroying the peer review infrastructure. He emphasized that it was important to have frank, open discussions on the issue and to determine how best to move forward to serve science.

Dr. Zerhouni stated that while NIH could not stay away from the issue and had to develop a policy, NIH also has created an adaptable process for change. NIH is looking for directions for further action and for ways to deal with the cultural change necessary to make the policy successful. He elaborated on the three goals of the policy, stating that the archive not only needs to enhance access to the literature, but also must be integrated with other tools, such as GenBank and the protein databases. Such integration of resources is essential for scientists to answer new questions without any artificial barriers in the way. These capabilities will enable researchers to see the process by which other scientists have reached their conclusions, accelerating discovery and creating tremendous value for NIH. Dr. Zerhouni also emphasized the value of NIH understanding where research investments have been made and where progress has occurred.

Finally, he emphasized the importance of providing access to the results of NIH supported research to the public at large.

Dr. Lindberg commented that there had been a discussion by the group about the value of journals providing the final version of an article to PMC. Dr. Sobel reiterated his concern about the quality of author provided manuscripts and the added value publishers provide in fact checking and other quality control measures. Dr. Zerhouni noted that while accuracy is important, being able to search and link manuscripts to other resources is also important. Ms. Terry commented that this provides a greater reason to make information available right away and to mandate compliance with the policy. Dr. Zerhouni responded that he believed it was important to start with flexibility in the Policy. Ms. Wolpert stated that viewing the NIH Policy as building not just an archive, but a tool that enables scientists to move across boundaries is very valuable. Dr. Zerhouni emphasized again the importance of developing these tools and connecting to other resources, including those in the private sector. This network effect will be extremely productive, but NIH recognizes that the Policy may require some adjustment over time to accomplish this interconnectivity.

Mr. Williams asked about whether NIH believes what it is doing will satisfy Congress. Dr. Zerhouni responded that Congress is interested in seeing the clinical impact of the results of NIH research and believes that better public access to research funded by the taxpayer is needed.

Dr. Thomas stated that Nature's policy supports self-archiving after 6 months, and they have informed NIH authors that they will upload their articles for them. She believes that it is important to get the message to Congress that a March reporting date on progress in achieving the goals of the Policy is unrealistic. Dr. Zerhouni responded that it is important to come up with a reasonable path for accomplishing the Policy's goals, that he is looking to the Working Group to help define what is reasonable and fair, and that NIH is willing to change the Policy as necessary. Dr. Thomas reiterated that in today's environment she believes what NIH is doing may be a type of publication. She inquired about what happens to material once it is in PMC. She noted that in order to determine the new value-added provided by publishers like Nature, they need to know what PMC will offer as the basic service to the public. She described the various roles a publisher plays in the journal publication process and raised the question of whether or not branding articles in PMC would be useful or not. Dr. Zerhouni replied that publishers need to view what NIH is doing as stimulating new value-added products, rather than NIH moving into the publishing business.

VII. NIH OUTREACH STRATEGY TO DATE

Dr. Timothy Hays described NIH's efforts to publicize the Public Access Policy. Outreach has taken different forms including such things as articles in the media; talks by NIH staff in a wide variety of venues, such as scientific meetings; and direct communication with affected NIH staff and investigators. NIH has held meetings for staff, sent emails to all investigators receiving NIH support, and developed an extensive website containing a variety of information on the Policy. The website is regularly updated and future emails will be sent to investigators to remind them of the Policy.

Ms. Terry stated that investigators would be more responsive if the emails came from their program officers, rather than a broadcast email. Dr. Hays agreed with her suggestion, but replied that NIH is not currently in a position to develop email lists of that type, and that it has taken considerable effort to assemble a reliable, comprehensive email list of all investigators.

Dr. Ward commented that there is a need to evaluate awareness of the policy versus compliance with the policy. Only in that way will it be possible to determine if people are not participating because they are unaware of the Policy or are aware, but have decided not to participate.

Dr. Thomas said that *Nature* would be willing to provide links to the NIH Public Access Website and perhaps other publishers would do so as well. She asked whether NIH would be doing a survey to determine why people were not participating. Dr. Ruiz Bravo indicated that it is difficult to undertake surveys in the federal government due to the clearances required, but NIH would look into the possibility.

VIII. RELATED POLICIES OF OTHER RESEARCH FUNDERS

Dr. Lipman discussed the recent policy statements by the Wellcome Trust and the Research Councils, UK (RCUK) concerning open access to published results of research they fund. Copies of the statements were distributed. The Wellcome Trust has announced that beginning October 1, 2005, all papers from new research projects must be deposited in PubMed Central or UK PubMed Central after its formation within six months of publication. The RCUK will require for all grants awarded after October 1, 2005, that subject to copyright and licensing arrangements, a copy of any published journal article or conference proceeding resulting from that research should be deposited in an appropriate e-print repository. Dr. Lipman indicated that about 10,000 articles a year are produced from research funded by Wellcome Trust and RCUK. Dr. Sobel made the observation that the RCUK policy stated that any repository was adequate and did not require deposit in a central repository. Dr. Thomas commented that currently the UK does not have a central repository and Dr. Lipman noted that once a central repository is established, there may be a requirement to deposit articles there as well as any other repository. It also was noted that the NIH Policy encourages deposition of manuscripts in multiple repositories, in addition to PMC.

IX. WORKING GROUP DISCUSSION AND DEVELOPMENT OF ANY RECOMMENDATIONS TO NLM BOARD OF REGENTS

Dr. Detre asked each member of the Working Group, as well as the senior NIH staff participating in the meeting to provide comments and make any recommendations they believed the Working Group should make to the NLM Board of Regents.

Dr. Sobel recommended that the Working Group should facilitate discussion to change the NIH Public Access Policy so that the publisher's final version, rather than the author's manuscript would be deposited.

Ms. Thibodeau agreed that it would be advantageous to get the publisher's version. She also suggested that NIH work more closely with libraries to do outreach at the local level.

Dr. Kamlet recommended that NIH get a better understanding of what is happening in response to the Policy and why, including perhaps conducting a survey. He suggested that NIH look for ways to make the Policy a "win-win" proposition.

Dr. King also agreed that it would be beneficial to work with publishers to increase submissions.

Dr. Koerner stated support for having the published version submitted to PMC. He advocated setting a realistic timeframe for working with the publishers to accomplish this. He also commented that providing public access to material written for the public was a separate topic requiring more discussion.

Dr. Drazen recommended that the journal of record be the primary place for users to view articles deposited in PMC and that PMC be used strictly for searching purposes. He was concerned about the impact on the branding of the journal if articles were viewed on the PMC website.

Ms. Wolpert recommended that the Working Group create a timeline and determine what needs to happen at specific dates. She emphasized that universities need to have a role in implementing the Policy.

Mr. Tykeson suggested that NIH make a distinction between grants previously funded and new grants. He proposed that on new grants NIH require the publisher to provide the final manuscript to PubMed Central within a reasonable timeframe (to be established) from the publish date, and that this policy be effective on a date certain giving adequate time for proper notice to all. All grants made before the effective date of this policy would be given a grace period, possibly as much as one year to comply. He emphasized that NIH should move forward expeditiously with the policy given that access to PMC has the capability to make scientific advances occur faster and result in improvement in health care.

Dr. Marcum stated that it is important to assist NLM in meeting its mission in the new digital environment. She suggested that a few publishers work with PMC to get more experience in how to facilitate the submission process.

Dr. Lipman reiterated the importance of determining the difference between knowing about the Policy and participating in it. Once that difference is known, NIH will have a better understanding of what needs to be done to improve participation.

Dr. Ward stated that NIH may ultimately need to mandate the Policy to get adequate participation. The Working Group needs to define what is meant by success and will likely need to conduct a survey to get needed data.

Ms. Terry stated that publishers should not underestimate the ability of the public to understand and make use of the biomedical literature. She recommended that NIH develop a comprehensive plan for adoption of the Policy and emphasize that the Working Group should not be content with waiting to see if the Policy needs to be changed given that access to these resources can accelerate scientific advances upon which people's health depends.

Mr. Williams recommended that ways be found to provide incentives so people will submit their manuscripts, including potentially having some mechanisms that involve the program officers. He stated that there is a need to collect data to show that the Public Access Policy has made a difference. He also believes that the delay period should be set at six months.

Dr. Ruiz Bravo reiterated the need to understand why submission numbers are low; whether there is lack of knowledge about the policy or lack of compliance. If it is lack of compliance, what are the causes?

Dr. Thomas agreed that a comprehensive plan, along with articulation of results that all can agree upon is needed. She recommended a multi-faceted outreach campaign that could be effective using the different stakeholders on the Working Group. She emphasized the need for a single message that is consistent and that all participants could use.

Mr. Nairn commented that he is not sure what NIH's ultimate goals are for the Policy. Initially it was described as creating an archive, and now it is described as developing a tool. He believes that the more NIH provides such tools, the more NIH becomes a publisher and that NIH is interweaving support for scientific advancement with public access. He recommended that NIH provide a clear strategic view of what is being built and that tactically there is a problem dealing with authors rather than publishers.

Ms. Humphreys emphasized the value of getting data on compliance versus awareness.

Dr. Detre stated that Congress clearly directed NIH to find a way to improve the sharing of the results of NIH funded research and recommended that the Working Group come up with unambiguous guidelines for accomplishing this goal so that all stakeholders understand what is being proposed.

Dr. Lindberg recommended that NIH move ahead with working with the publishers to get the final versions of published manuscripts.

Dr. Detre concluded the meeting by stating that the next meeting will be scheduled for September and members will be contacted concerning their availability. He suggested that if members have issues they would like to see raised at the next meeting, they should send them to him. The meeting was adjourned at 3:00 pm.