NIH Public Access Working Group of the NLM Board of Regents
Meeting Summary
April 10, 2006

Members Present:
Dr. Thomas Detre, Chair; Dr. Deanna Marcum, Dr. Susan Buchanan, Dr. Jeffrey Drazen, Dr. Samuel Kaplan, Dr. Mark Kamlet, Dr. T.J. Koerner, Mr. Brian Nairn, Dr. Mark Sobel, Ms. Patricia Thibodeau, Mr. Donald Tykeson, Dr. Gary Ward, Ms. Ann Wolpert

Members Absent:
Dr. Harvey Fineberg, Mr. Michael Stern, Ms. Sharon Terry, Dr. Annette Thomas

Staff:
Dr. Donald A.B. Lindberg, Ms. Betsy L. Humphreys, Dr. Donald W. King, Dr. David Lipman, Dr. James Ostell, Mr. Kent Smith, Mr. Ed Sequeira, Mr. Robert Mehnert, Dr. Neil Thakur, Dr. Dennis Benson, Ms. Becky Lyon, Mr. Sheldon Kotzin, Dr. Elliot Siegel, Ms. Rebecca Wilson, Dr. Bart Trawick, Dr. Kerry Zbicz

Members of the Public:
Mr. Robert Harington, Mr. Michael Weston, Ms. Laura Brockway, Ms. Janet Coleman, Ms. Heather Joseph, Mr. Eric Massana, Dr. Martin Frank, Mr. Jerry Sheehan

Dr. Lindberg welcomed the members and thanked them for their contributions. He noted that there was recent correspondence pertaining to the Working Group's work and he turned the meeting over to Dr. Detre.

Dr. Detre also welcomed the Working Group. He noted that the members had previously received copies of two letters both dated March 31, 2006—from NIH Director Dr. Elias Zerhouni and jointly from Elsevier CEO Brian Nairn and Dr. Mark E. Sobel, Executive Officer of the American Society for Investigative Pathology. [Both letters are in Appendix A.] Dr. Detre said that Dr. Zerhouni was recently queried by Representative Ernest J. Istook (R-OK) about progress in Public Access. Dr. Istook asked for a report from the NIH on the subject. Dr. Detre then asked each member of the Working Group to comment on developments since the last meeting on November 15, 2005.

Dr. Samuel Kaplan (who has succeeded Mr. James Williams as an ex-officio member of the Working Group) introduced himself: he is Professor of Microbiology and Molecular Genetics at the University of Texas Health Science Center (Houston), Chair of the Publications Board for the American Society for Microbiology (ASM), and chair of the Advisory Committee for PubMed Central. Dr. Kaplan asked about the current state of the Public Access initiative. Dr. Detre, summarizing their last meeting, said there was no disagreement that making taxpayer-supported publications available to the public was the "right idea." It was agreed that the best way would be for the publisher to submit to NLM the final version, signed off on by the author. This would ensure quality and accuracy. There was discussion about how much time should be allowed for this to happen: some

felt that six months was proper, others felt that one year would be better, especially for publications that appear less frequently. The question now is where should we go from here? How do we develop a consensus that makes sense and that the Congress would accept?

Dr. Kaplan reported to the Working Group about the experience with the 11 journals of the American Society for Microbiology. For the last two years, all of their material has been deposited in PubMed Central with a six month embargo period, except for a 12 month embargo for two review journals. It has been the experience of the Society that the number of personal subscriptions has dropped, although this was anticipated and has nothing to do with deposit in PubMed Central (PMC). However, there has not been any decline in income from the journals, and the number of accesses through PMC has gone up dramatically, while access at the ASM's own journal sites has continued to grow. The bottom line is that there has not been any deleterious effect of making the articles available.

Mr. Tykeson asked about how easy (or not) it is for authors and publishers to comply and submit material for PubMed Central. Couldn't this be done automatically by NIH when NIH grants officers receive the published material as part of grantees' regular reports? Dr. Sobel said that after a paper has been accepted for publication, there is often back and forth communication with the author concerning editing, formatting, checking data, and ironing out inconsistencies. After that process, the paper ends up online, in print, or both. It would be a tremendous administrative burden, he said, for those responsible for handling NIH grant applications to be responsible for ensuring the submission of publications to PubMed Central. Authors who submit their manuscripts to PubMed Central automatically have them included as part of their reports to NIH, taking the onus off them in their communications with NIH. It was clarified that the system being suggested is only for manuscripts submitted and accepted for publication, not unpublished interim or final reports to the NIH. Dr. Ward emphasized that the system for submitting journal articles to PubMed Central was neither onerous nor burdensome—participating publishers have worked with PMC and the posting process is "seamless." In the case of publishers who don't participate, it took him less than 10 minutes to submit one of his own articles to PMC, and then 10 minutes to proof the final.

Dr. Kamlet said that several publishers have complained to him that manuscripts they sent to PubMed Central on behalf of authors were changed. Dr. Lipman said the manuscript sent to PubMed Central by publishers is generally the author's word processing document. NLM contractors tag this manuscript into an archival format and then it is sent back to the author for review. Because word processing documents are very "non-uniform" in terms of format, a table might be misaligned, for example, and inadvertently change the content. That is why the author must proof and approve the XML version of an author's manuscript before it is released into PMC. An automatic conversion is made for the final published articles sent by publishers that participate in PMC and send tagged documents to the NLM. Dr. Lipman also said that they are trying to improve the "process of discovery" by tightly coupling the content of the papers in PMC with underlying factual databases.

In response to a question from Dr. Ward, Ms. Humphreys characterized Dr. Zerhouni's response to Mr. Istook's queries about the effectiveness of the current voluntary policy: that it appeared that the voluntary approach was not working and that NIH would look at the policy in that light. Dr. Detre expanded on Dr. Zerhouni's reply to Mr. Istook by saying that there are different views about the six months versus one year and that the NIH Director did not commit himself on this matter.

Ms. Wolpert said that at MIT they found that their principal investigators dealt with more than 30 publishers in 2005. It can be a complicated matter for a PI to deal with multiple publisher agreements and processes.

Dr. Drazen said that he would like to get a sense of the group for the answer to three questions: (1) Should the policy be voluntary or required? (2) If it is required, what time frame should be specified? (3) If it is required, by what means or entity do the manuscripts become publicly available? He said that as for his organization (*New England Journal of Medicine*), the policy should be mandatory, that six months is the right time frame, and that the published version should be what is made available. They do have a question as to whether the article should be made available in an "expanded Medline/PubMed" or in PubMed Central. Dr. Drazen added that only about 30% of the research articles they publish receive NIH funding. (*NEJM* also publishes articles that do not report original research.) He suggested that ways be found to encourage publishers to deposit all articles, not just those that are NIH-funded. Dr. Neil Thakur, NIH Program Manager for Public Access, pointed out that one fundamental goal of the Public Access policy is to create an archive of NIH funded work for internal portfolio management purposes. This is achieved by requiring deposit in PubMed Central.

There was a discussion of the value of "branding" a journal and whether archiving articles in PubMed Central would jeopardize a journal's brand. Dr. Kaplan said that before deciding to deposit articles in PMC, his society considered the issue of branding very carefully. The result of their depositing articles "has been wonderful—there has been absolutely no problem…Everyone knows who is publishing these papers—the ASM." Initially they had a one-year system, but they moved it to six months and are now even considering four months.

Dr. Sobel said that there are several constituencies to be considered—the public, the NIH, and the scientific community. We should look at the various mechanisms that exist for the various constituencies, prominently among them personal and institutional subscriptions. Every journal also has a "pay for view" mechanism. Most publishers respond to individuals who request an article. Everything in the health literature thus is available. The questions are how cumbersome is the process and what will it cost?

Dr. Ward commented that for educators and researchers, these mechanisms are not that useful because it is necessary to be able to browse the literature. Although in principle, material is available, in terms of "how we do our work," it isn't. As to the issue of time, he said that his journal offers content after two months and subscriptions continue to rise.

Dr. Lipman put the question in terms of "how much of an obstacle does there need to be for people to stop looking for something?" He said his impression is that even a small obstacle will prevent people from looking. Dr. Sobel reminded the group that in fact Medline/PubMed has abstracts and makes possible successful searching of the literature in some detail.

Mr. Nairn said that it has become clear that the issue is between NIH and the publishers. We are saying that we want the final published version of an article—in which the publisher has invested much money. The real question is how we get the final version for the NIH while at the same time protecting the business model of the peer-review process. There is "unease" as to whether we are giving the publisher's "intellectually aided" product to NIH to create something that then becomes a competitive product. Dr. Detre responded that the other side is that American citizens paid for the research through the NIH and that the journals that publish the results benefit from this investment. What would be a good synergy between the two to allow the results to be available to everyone within a reasonable period of time? Mr. Nairn said that there are two objectives. The first, public access to scientific materials, can be easily met and publishers will not stand in its way. The second issue is the concern of the publishers about the linking of materials and the value that is thus being added to the content supplied by publishers to NIH. This is the crux of the situation and forms the basis for his current discussions with the NIH.

Dr. Lipman presented to the Working Group statistics about compliance with the policy, announced in May 2005, requesting that authors submit final articles for archiving in PubMed Central. Since that time there has been virtually no change (2.3%) in the number of principal investigators submitting manuscripts. A higher percentage of authors, however, are reviewing their data before actual public release. The turn-around time for tagging manuscripts is now two weeks or less. The submission pipeline—software and workflow—that has been created by NIH is working very well, within budget, and capable of being scaled up. Dr. Lipman described in some detail how the process works, from the acceptance of a manuscript by a journal to its final appearance in PMC. He also noted that some publishers have expressed interest in using the regular PubMed Central process (submission of tagged electronic copies of the final published versions of articles) only for those articles that have NIH funding. This was termed the "Partial PMC" approach for the rest of the Working Group's discussions. If any publishers elect this approach for submission of NIH-funded articles, NLM would be able to do an automatic conversion for PMC.

Dr. Lipman gave statistics on the number of articles that NLM has removed from PMC in response to notifications from publishers that NIH-funded investigators had submitted them in violation of pre-existing copyright agreements. He also described how NLM makes corrections—retractions, errata, etc. both to the citations in PubMed and full-text articles in PubMed Central. NLM also can provide statistics to the publishers and authors about how often their publications are requested.

There was a general discussion about the percentage of articles appearing in journals authored or co-authored by NIH-supported scientists. The number varied by discipline

and the nature of the journal. This led to a discussion of embargo dates and the economic impact of public access on journals.

Ms. Wolpert summarized results from a survey by the Association of Learned and Professional Society Publishers. The 340 responses, primarily from academic institutions, found that the three most important factors to determine journals for cancellation were (1) the faculty no longer required them, (2) declining usage, and (3) cost. Availability of a journal via "public access" after some embargo period was far down on the list of factors.

Dr. Ward commented that, although faculty at major institutions do in most cases have access to the journal literature, faculty at many mid-level research universities do not have such access. Trying to do NIH-funded research where you don't have access to those materials is difficult. Ms. Humphreys agreed, and said that the Congress has also noted the problem of disparity in how research dollars are distributed and there is concern about how to build up the research infrastructure in states that do not receive much NIH funding. Young people from disadvantaged groups might be more likely to go into science if they were at institutions that had improved access to the biomedical literature.

Dr. Marcum said that we must guard against trying to use a new initiative to solve "all existing problems." We should focus on the main issues that were put before the Working Group. Dr. Detre said that the charge is clear. Because what is being proposed is something new, it is reasonable to express doubts and concern about possible consequences of our actions. Mr. Tykeson said that, based on Dr. Zerhouni's comments and the expressions of the Congress, the efforts of this Working Group will be irrelevant unless there is "a decision for mandatory submission to PMC." He suggested that the Working Group should agree on mandatory submission within six months and then focus on the details of how this could be attained.

Dr. Kaplan said that his ASM journals have a vetting process to deal with "dual use," that is, information in a journal article that might be used for terrorist purposes. Another issue is that of assigning authorship—in some cases between the time a manuscript is accepted and when it finally appears in the journal there are changes in the authorship. There is also the issue of the "corresponding author" and the NIH-funded author; they are not necessarily the same person. These are issues that would have to be attended to. They argue for using the publisher's final version in PMC.

Dr. Neil Thakur gave the Working Group an overview of recent meetings NIH has had on the subject of Public Access with three interest groups. There have been several meetings with Elsevier to discuss the batch submission of 100% of their NIH-funded author manuscripts. There have been discussions with the Coalition of Nonprofit Publishers about receiving 100% of their NIH-funded articles—in this case we are talking about the final copy-edited version. This is what was referred to earlier as "Partial PMC." There has also been a meeting of NIH with the Alliance for Taxpayer Access to discuss ways the public uses research and how the Public Access initiative can help. In response to a question from Dr. Koerner about who is participating in the discussions with NIH, Dr. Thakur said that "we are willing to talk to anyone who wants to talk with us." There are

specific negotiations with specific groups based on how they define themselves. PubMed Central and the submission infrastructure form the bedrock of the discussions on how the goals can best be accomplished.

The members of the Working Group received before the meeting information about a recent report about NIH Author Postings that was prepared on behalf of the Publishing Research Consortium. PowerPoint slides based on the report's findings were distributed at the meeting. Mr. Brian Nairn briefly described how the survey of authors was done and he went through some of the highlights of the report. He noted that the results basically corroborate information presented by Ms. Thibodeau at the November meeting.

Dr. Susan Buchanan said that it is her experience that intramural scientists in her Institute (NIDDK) are only vaguely aware of the issue and they don't understand the benefits of depositing manuscripts in PMC and how easy it is to do. If we have a mandatory Public Access process, there will definitely need to be more outreach and education.

Dr. Sobel predicted that since we are now 11 months after the policy went into effect, we will see a "big bump" in submissions, probably by July. He raised the question of what kind of response NIH grant administrators got in their dealing with grantees about the Public Access issue. Dr. Ward said that as a grantee he gets lots of e-mail on the policy from NIH.

Ms. Wolpert raised the "chilling" effect that copyright has on authors and their willingness to participate in Public Access. She presented a report to the Working Group on what they have learned in trying to facilitate depositing of peer-reviewed manuscripts by MIT faculty. NIH accounts for as much as a third of sponsored research at MIT. The question was how to extend the existing framework that supports students, faculty, and scientists at MIT to be supportive of the Public Access policy. They discovered, first, that their scientists frequently didn't have the final peer-reviewed version of the manuscript. Second, the scientists didn't have legal agreements with their publishers that would enable them to deposit a manuscript in PMC. Ms. Wolpert said that, with the involvement of the MIT attorneys, they drafted an agreement that would deal with copyright issues and permit participation. The agreement received the endorsement of department heads and the MIT faculty and will be implemented over the next several months. The faculty has asked that a model copyright agreement, covering monographs and journal articles, be put on the Web so that publishers would know what the MIT policy is. She added that in her experience she has never found that the availability of work in a public repository (after an embargo period) has had any impact whatsoever on decisions about subscriptions. Decisions are entirely based on faculty needs, utilization, and costs.

Dr. Detre asked the Working Group to consider the three questions raised earlier by Dr. Drazen.

(1) Should depositing articles in PMC be voluntary or required for NIH grantees?
(2) What should be the time requirement?
(3) What should be the form of the submitted article?

The majority of members confirmed the opinions expressed at the previous (November 15, 2005) meeting of the working group: (1) the policy should be mandatory; (2) submission should occur within six months (with flexibility to 12 months in the case of journals that publish quarterly or less frequently); and (3) the final manuscripts as published should be the favored form. A minority favors a 12 month submission deadline and submission of the author's final manuscript rather than the final published form.

Mr. Nairn said that the best way to increase participation in the short term by the publishers is for NIH to engage with them about alleviating the publishers' concerns about what is going to be done with the archive (the links and the "value added") and then encourage them to engage in bulk submission of articles.

Dr. Thakur noted that whatever recommendations are adopted will not be implemented overnight by NIH. There will be a public education period, a phase-in period, and there may be a period for comments. Mr. Tykeson said that there should be a timetable for implementation that gives all parties ample time to conform and that includes publicity so that the policy is known to all. We should also take one final look to ensure that the implementation process is as simple and easy as possible for the scientists and publishers.

Dr. Detre closed the meeting by saying that "the public good" that will result from these arrangements must be emphasized in announcements and descriptions of the policy recommendations. The public good is what originally prompted the Congress to ask that the matter be looked at and recommendations made.