

Public Access to NIH-Funded Research

Robert Steinbrook, M.D.

A new era for online public access to the biomedical literature is about to begin. As of May 2, the National Institutes of Health (NIH) has asked the investigators it funds to submit voluntarily to PubMed Central (www.pubmedcentral.nih.gov) an electronic copy of any scientific report, on acceptance for publication, and to specify when the article should become public¹ (see box). According to the NIH, "Posting for public accessibility through [PubMed Central] is requested and strongly encouraged as soon as possible (and within twelve months of the publisher's official date of final publication)."^{1,2}

Currently, about a third of the reports of recent NIH-funded research are publicly available in electronic form after a 12-month delay — but from a variety of repositories and in various formats, according to Dr. David Lipman, the director of the National Center for Biotechnology Information at the National Library of Medicine, where PubMed Central was developed and is operated. Thus, the centralized archive may become a leading electronic database of biomedical literature. Articles are available without charge to the user, and registration is not required. The NIH funds 212,000 researchers worldwide, and 5000 scientists are direct employees of the institutes. Each year, these researchers publish 60,000 to 65,000 articles, accounting for about 10 percent of the articles in the nearly 5000 journals indexed by PubMed. According to the NIH, "As the electronic article increasingly becomes the authoritative and most useful document for researchers, and as scientists are actually computing on the contents of these documents — the text itself as well as the associated data — the impermanence of the publishers' Web sites presents a substantial risk. Creating such an archive is a historical and necessary NIH responsibility."³

PubMed Central is linked to — but differs from

— PubMed, the National Library of Medicine's enormous database of citations and abstracts, which is also freely available. PubMed includes more than 15 million citations of articles in biomedicine dating back more than 50 years. It is the database most widely used by researchers, clinicians, and the public to locate articles and, in many instances, to link directly to the full text of articles at journal and other Web sites.

PubMed Central was launched in February 2000 with content from the *Proceedings of the National Academy of Sciences* and *Molecular Biology of the Cell*. Articles located by a search are automatically linked to other databases at the National Library of Medicine, such as GenBank and Online Mendelian Inheritance in Man. Currently, PubMed Central contains about 350,000 reports (mostly articles, but some letters) from about 180 journals. Within the year, with the addition of articles from back issues of these journals, articles from new journals, and the public-access articles, it is expected to contain between 700,000 and 800,000 reports. Articles are linked to PubMed and to Web sites maintained by publishers.

The NIH was spurred to develop its public-access policy by congressional pressure mounted as part of the budget process and by groups, such as the Public Library of Science, advocating "open access" to the biomedical literature. According to Dr. Elias Zerhouni, director of the NIH, the goal is to make "a change in the landscape of how scientific information is made available to the public while preserving the viability of the peer-review process."

The NIH's initial proposal would have required that scientific reports supported by the institutes be made available at PubMed Central no later than six months after they appeared in a journal.⁴ Subsequently, the agency reviewed more than 6000 comments and revised the policy "to provide flexibility to ensure maximum participation," Zerhouni said in February. The NIH, which has an annual budget

Dr. Steinbrook is a national correspondent for the *Journal*.

Implementation of the Public-Access Policy

As of May 2, 2005, NIH-funded investigators are asked to submit voluntarily to PubMed Central (www.pubmedcentral.nih.gov) an electronic version of the author's final manuscript when the article is accepted for publication. This version is defined as "the final version accepted for the journal publication, and includ[ing] all modifications from the publishing peer review process." Included are publications resulting from current projects that are funded in whole or in part by the NIH, as well as reports on previous NIH-supported research that are accepted for publication on or after May 2.

The policy applies to all research grants, career-development awards, cooperative agreements, contracts, and national research service awards, as well as NIH intramural research studies. It does not apply to book chapters, editorials, re-

view articles, or conference proceedings.

Manuscripts are to be submitted in the usual word-processing or PDF formats through a secure Web-based system. There are procedures to ensure that submissions are consistent with copyright assignments and agreements and that the journals have been notified of the submission.

At the time of submission, the responsible author will specify when the manuscript is to become publicly accessible through PubMed Central; no article will become accessible until after it is published.

The National Library of Medicine will use a standard digital archival format to store manuscripts. After manuscripts have been converted to this format, the responsible author will be sent an electronic copy. No manuscript will be released until the author has verified its accuracy.

If the publisher provides its final version of the article, this version will supersede the author's final version. When publishers transmit manuscripts, the National Library of Medicine will ask the responsible author to sign off on the transfer and verify key information, such as the accuracy of the paper and the release date. The author's final manuscript will still be available at PubMed Central, through a link from the publisher's final version. If the publisher agrees, public access to the publisher's final version can occur before the time originally specified by the author.

As the NIH gains experience with the new process, the policy will be refined. An advisory committee, the NIH Public Access Working Group of the Board of Regents of the National Library of Medicine, is being established.

of about \$28 billion, estimates that the annual incremental costs of the public-access policy will be \$2 million to \$4 million. In addition to improving public access and creating a stable electronic archive of publications from NIH-funded research, the NIH plans to use the database to improve the management of its research investment, to monitor scientific productivity, and to help set research priorities.

In the near future, however, the initiative may have other effects. First, it may confuse investigators, many of whom may be unaware of the policy or what it does and does not entail. For example, an author's final manuscript — the version the NIH has requested — is the manuscript as accepted by the journal, including all changes resulting from peer review. If the publisher chooses to furnish its final version of the manuscript — the version that also includes editing and other changes that occur after peer review — it will supersede the author's final version.

Second, the initiative may continue to rile some journal editors and publishers. Some view aspects of the policy — such as its strong encouragement of ensuring public accessibility of manuscripts "as soon as possible" — as potential threats to the integrity of their content, their Web sites, and their revenue sources. Although the policy states that "the

author will specify the timing of the posting of his or her final manuscript," many journals seek to maintain a 6-month or 12-month delay between their publication of an NIH-funded study and its availability through PubMed Central. Some have announced their intention of transferring articles directly to PubMed Central on behalf of their authors, with instructions about when the articles should be made public.

A specific concern is the effect of posting a version of a manuscript that may include uncorrected content errors or conclusions that are later revised, particularly if the article has implications for patient care. In response, the National Library of Medicine has said that PubMed Central will be able to accommodate corrections of content errors and other necessary revisions to manuscripts.² In practice, this may be difficult to do on a consistent basis, particularly when timelines are tight or if PubMed Central is not notified of the necessary changes.

The NIH has repeatedly stated that the policy is a request rather than a requirement and that it will neither monitor the compliance of individual investigators nor take information about compliance into account in making future decisions about funding.¹ According to Zerhouni, "We trust our scientists. We believe they will do what is right for the public who funds them." Still, investigators may be

caught between the NIH policy and the policies of journals in which they seek to publish their work, and journals and publishers may be caught between their support for the public health mission of the NIH and their own self-interest.

The NIH initiative may also encourage other governments and private funding organizations to consider public-access policies. For example, the Wellcome Trust and the National Library of Medicine are discussing the establishment of a counterpart to PubMed Central in the United Kingdom.⁵ Such a repository would mirror the data held in PubMed Central and also provide the flexibility to add additional publications and content.

As the public-access policy takes effect, there are high expectations for quick movement toward timely availability of all publications from NIH-sup-

ported research. PubMed Central, however, could soon receive 5000 papers a month, or only a few hundred. It should rapidly become obvious whether the policy is working as the NIH — and Congress — intended.

1. Department of Health and Human Services, National Institutes of Health. Policy on enhancing public access to archived publications resulting from NIH-funded research. *Fed Regist* 2005; 70(26):6891-900.
2. National Institutes of Health. Final NIH public access policy implementation. March 15, 2005. (Accessed April 8, 2005, at <http://www.nih.gov/about/publicaccess/Finalpublicaccessimplementation031505.htm>.)
3. *Idem*. Public access questions and answers. February 24, 2005. (Accessed April 8, 2005, at <http://www.nih.gov/about/publicaccess/index.htm>.)
4. Department of Health and Human Services, National Institutes of Health. Enhanced public access to National Institutes of Health (NIH) research information. *Fed Regist* 2004;69(180):56074.
5. Terry R. Funding the way to open access. *PLoS Biol* 2005;3:364-6.

Orchestration of Iron Homeostasis

Robert E. Fleming, M.D., and Bruce R. Bacon, M.D.

Related article, page 1769

The number of newly identified genes participating in the regulation of iron homeostasis has increased at a remarkable pace. The characterization of these genes has led investigators to challenge previous models of the regulation of iron homeostasis in health and its dysregulation in disease. There is now substantial evidence that the liver plays a central role in determining how much iron is absorbed from the gut and in influencing the release of iron from sites of storage. The discovery of the iron regulatory hormone hepcidin has provided a cohesive theory to explain the pathophysiology of such common disorders as hereditary hemochromatosis and the anemia of inflammation (also known as the anemia of chronic disease). The most important cellular targets for hepcidin appear to be the villus enterocyte, the reticuloendothelial macrophage, and the hepatocyte (see diagram).

There are no substantial physiologic mechanisms that regulate iron loss. Accordingly, iron ho-

meostasis is dependent on regulatory feedback between body iron needs and intestinal iron absorption. Several factors have been found to influence the rate of iron absorption, including the body's iron stores, the level of erythropoietic activity in bone marrow, the blood hemoglobin concentration, the blood oxygen content, and the presence or absence of inflammatory cytokines.¹ More than one of these factors may act simultaneously, and some are interrelated. Intestinal iron absorption increases with decreased iron stores, increased erythropoietic activity, anemia, or hypoxemia. Conversely, intestinal iron absorption decreases in the presence of inflammation — a process that contributes to the anemia of inflammation. Excess iron absorption relative to body iron stores is the hallmark of hereditary hemochromatosis.²

Nearly all absorption of dietary iron occurs in the duodenum. Several steps are involved, including the reduction of iron to a ferrous state, apical uptake, intracellular storage or transcellular trafficking, and basolateral release. Molecular participants in each of these processes have been identified³ (see diagram). The reduction of iron from the ferric to the ferrous state occurs at the enterocyte

Dr. Fleming is an associate professor of pediatrics and biochemistry and molecular biology and Dr. Bacon a professor of internal medicine at the Saint Louis University School of Medicine, St. Louis.