



NIH Releases Final Policy on Centralized Database of Human Genetic Data

The National Institutes of Health announced on Tuesday the final version of a policy governing a powerful, central repository of human genetic data that the agency will establish for the study of diseases.

Among other provisions, the policy will give scientists who contribute data to the repository exclusive rights for up to 12 months to publish scholarly analyses based on the data.

In other respects, the final policy, published in the *Federal Register*, is similar to a draft version published last year for comment (*The Chronicle*, August 21, 2006). The agency said it received nearly 200 comments, a large response reflecting controversy over some aspects of the new database, including privacy protections and controls over publishing and intellectual property.

The NIH and biomedical researchers are viewing the database as a promising tool that will help bring about the next phase in the development of medicine: new understanding of the contribution of genes to many common diseases. Scientists have already begun such research using smaller databases. Such work aims to correlate, for example, conditions like heart disease and diabetes with variations in particular genes.

The success of such studies, called genome-wide association studies, depends on having data from large-enough populations to yield statistically significant results. The NIH's data repository will allow the pooling of smaller databases containing similar information. Scientists financed by the agency to conduct such studies will be required to contribute their data to the repository.

Many of the comments received by the NIH questioned whether the agency could adequately protect the privacy of patients whose genetic code was catalogued in the database. The NIH will require that researchers strip the data of details that could identify individuals. (The NIH will make public summary descriptions about data on particular diseases.)

The agency will also require scientists who want to study the database to agree not to publicly distribute it. Privacy advocates have worried that health insurers, for example, might obtain information from the database that identified individuals as having a genetic susceptibility to certain diseases and deny them coverage as a result.

In addition to establishing other privacy protections, the policy urges investigators to ask donors of DNA samples to sign a so-called certificate of confidentiality as well as the required informed-consent forms. The certificate confers extra protections that would allow the agency to resist legal efforts to compel disclosure of data that would identify individual donors.

To ensure informed consent, investigators might need to go back to patients who had previously donated tissue samples for research and verify that they agreed to the intended genetic studies, the policy notes. The NIH said it may consider giving money "or other resources" to help researchers with this extra work.

As for publishing, the NIH had originally proposed embargoes of up to nine months on the publications of research results based on the data by anyone other than the investigators who



contributed the information, but in the final version extended the limit to 12 months. Individual NIH institutes may choose to set shorter embargoes. Some commenters had complained that a nine-month embargo could run out before the original investigators' publications appeared and could lead scientists to rush into print hasty, premature analyses.

The final policy also clarified the original proposal to specify that publication encompasses forms of public dissemination besides peer-reviewed journals, including meeting abstracts, oral presentations, and Web sites and blogs.

As for intellectual property, the NIH's policy says that it will expect users of the database not to use patents to block other researchers from using that data.

The policy will take effect in January.