



NIH proposes centralized database of human genetic data to advance health research

(By SAM KEAN; reprinted from The Chronicle of Higher Education, August 31, 2006.)

To hasten the flow of genetic information among scientists, the National Institutes of Health plans to develop a repository that would collect and collate human genetic data, but in a way that would protect the privacy of people who participate in genetic studies.

According to a [notice](#) published in the *Federal Register* on Wednesday, the results of all NIH-sponsored research in "genome-wide association studies" would be deposited in the genetic database. Such studies analyze the genetic differences between people with a specific disease and healthy individuals.

Approved scientists would be allowed to use the centralized repository to find relevant health information about anonymous human subjects who share a certain trait -- say, heart disease or diabetes. The researchers would then be able to search those people's DNA for similarities in order to tease out potential genetic contributions to those conditions.

By bringing together many records for comparison, the depository should facilitate such research, which, though laborious, is a necessary step in developing one of the more promising areas of biomedicine: genome-based medicine. A spokesman for the NIH said that up to 20 projects that would contribute to the database are being planned or are already under way.

The idea of a single, central repository raises ethical and privacy concerns, however, especially if genetic records could be linked with individual patients. An insurance company, for instance, could use such information to discriminate against people with a propensity to develop a certain genetic disease. And while data is supposed to be submitted to the NIH repository in a form that protects patients' privacy, bioethicists suggest it would be nearly impossible to make the database risk-free.

"It's very, very, very difficult to provide ironclad promises to anyone about the privacy of genetic information," said Glenn McGee, director of the Alden March Bioethics Institute at Albany Medical College. "Anyone who tells you otherwise is selling you something."

Mr. McGee acknowledged the scientific value of such a database, however, saying "the notion of creating a repository was unavoidable."

He added that, despite the promise of genome medicine, most people who support the database weren't necessarily looking for immediate breakthroughs. "A long chain of modest successes would be just as valuable," he said, calling this the "NASA effect": "They find a microwave oven when they really meant to go to the moon."

The NIH points out that most of the patients whose information could end up in the central database would presumably already have given consent to participate in a smaller study at a university or laboratory. But before data from those smaller studies end up in the central repository, the NIH said it would seek "re-consent."

Furthermore, in the *Federal Register* notice, the NIH said it would require all genetic data to be submitted "without identifiable information." Submissions, it added, "should be accompanied by a written certification stating that the identities of research participants will not be disclosed" to the repository or to scientists doing secondary studies "without appropriate institutional approvals."

The genetic information would be available on two levels. "Basic descriptive information" about each genome-wide association study would be available to the public, but the health and genetic data and some pre-analysis would be accessible only to researchers cleared by the NIH Data Access Committee. The NIH would also set up mechanisms to ensure that scientists who deposit results into the database before publication receive credit for their findings. Once results have been posted, any scientist could view them; but for a certain "grace period," only those who submitted the data could publish papers about it.

The agency is seeking public comments on various aspects of the repository, including sharing data from it and questions about privacy issues and intellectual property rights. Comments will be accepted until October 31, and a town-hall meeting will follow in early December.