

Safeguards in the use of DNA databanks in genomic research

Sara Taub, MBe, Karine Morin, LLM, Robert M. Sade, MD, and Monique A. Spillman, MD, PhD for the Council on Ethical and Judicial Affairs of the American Medical Association

Genomic research, which uses DNA to identify the gene(s) responsible for complex diseases, relies upon large DNA databases to facilitate population-based research. These databases derive their power from integrating different kinds of information about large samples of individuals: genetic and clinical data and data on health, lifestyle, and environment. Using statistical analysis, the databases can be used to analyze correlations that may enable new therapeutic developments.

The use of DNA databanks in genomic research brings new challenges related to the scope of research as well as to the nature and use of the samples and information archived in such databases. This report identifies these issues and considers their relation to the current standard of informed consent in human research.

NEW CHALLENGES

Nature of DNA material

At the outset, several characteristics that are unique to DNA material (DNA samples or information derived from them) should be acknowledged. For instance, DNA material may include information about patients and their immediate biological family members that is entirely unknown to any of them—either because it has not been analyzed or because, in the absence of more advanced technology, it is not yet decipherable. Another characteristic specific to genetic information includes the fact that it may be used to forecast disease long before a person shows any symptom. Also, given the stability of the DNA molecule, the nucleotide sequence in samples is usually immutable. More importantly with regard to research, once collected and stored, a DNA sample can be duplicated almost indefinitely by polymerase chain reaction (PCR) and used in the future to answer questions that were not contem-

plated at the time the sample was obtained. Once a researcher has a subject's DNA material, the stored sample or the data derived from it can be used for any number of future research protocols.

These various features may raise special risks for individual research subjects whose DNA material is stored in databanks. The harms that could result to subjects from these risks are distinct from the physical harms that generally are associated with clinical research trials, such as the harm caused by the side effect of an experimental drug. Foreseeable harms from genomic research include insurance, employment, and education discrimination, social stigmatization, improper attempts to influence reproduction decisions, and distress caused by information regarding the statistical possibility of disease. These may be a source of distress not only for research subjects, but for their family members also.

Another new challenge that arises from genomic research is that some risks, including the possibility of stigmatization and discrimination, may extend beyond individuals to an entire identifiable community. This can happen when DNA data sets are constructed around relatively homogenous populations to increase the chances of detecting genetic variation within the < 0.1% deviation that exists between any two persons in the world. Any benefits that such population-based genomic research presents must be weighed against the possibility of these risks.

These considerations make it necessary to examine whether current standards that govern research can minimize the risks inherent to genomic research and sufficiently protect individual subjects as well as populations.

Uses of the data

In addition to acknowledging concerns that arise from the nature of the information archived in genomic databases, it is necessary to address intended uses of DNA material. Participating individuals and groups may have reservations about their information being utilized for certain types of research projects. Therefore, it is important that subjects have the opportunity to be informed about, evaluate, and consent to the goals of the intended research.

STRUCTURE OF CONSENT

Population-based genomic research raises the question of whether consent needs to be obtained solely from individuals or whether review by the target population is needed as well.

From the Council on Ethical and Judicial Affairs, American Medical Association, Chicago, Illinois.

Sara Taub, MBe, Senior Research Assistant, Council on Ethical and Judicial Affairs, American Medical Association, 515 N. State Street Chicago, IL 60610.

The Council on Ethical and Judicial Affairs of the American Medical Association formulates ethical policies for the medical profession through its interpretations of the AMA's Principles of Medical Ethics. The Council at the time this report was adopted consisted of Frank A. Riddick, Jr, MD (Chair), Leonard J. Morse, MD (Vice-Chair), Michael S. Goldrich, MD, Mark A. Levine, MD, John M. O'Bannon, III, MD, Priscilla Ray, MD, Naheed Rehman, MPH, MS, MA (student member), Robert Sade, MD, and Monique A. Spillman, MD, PhD (resident member). Staff to the Council at the time the report was adopted were Audiey Kao, MD, PhD (Vice President, Ethics Standards Group); Karine Morin, LLM (Secretary), Amy Bovi, MA, and Sara Taub, MBe.

DOI: 10.1097/01.GIM.0000144070.93743.3B

The concept of community review, also referred to as community consultation, may prove to be a successful method to minimize harm to certain groups and to help identify community support for population-based research endeavors. Examples of community consultation are the series of public meetings held in Oregon leading up to establishment of that state's Medicaid priorities¹ and the referendum held in Iceland before its adoption of presumed consent for genomic research.² When a community is opposed to the research, the study should not be conducted. When a proposal is met with support from the community, it nevertheless remains a requirement to obtain individual subjects' consent. Community consultation cannot be used as a substitute for informed consent. Instead, it should precede and complement the process, serving as an opportunity to begin educating members of the group from which research subjects will be drawn.

Informed consent

Although an imperfect safeguard, self-determination through informed consent has been considered an important mechanism to protect subjects from abuses in research. When the process is performed properly, it should prepare individuals to identify, understand, and consider the relevant risks and benefits that a research protocol presents. As a result of the new challenges that genomic research poses, additional safeguards may be necessary to address risks that arise from archived information and subsequent studies.

Consistent with the informed consent process in any type of human subjects research, subjects in genomic research should be informed of the following: (1) the purpose of the research; (2) the overall risks and benefits associated with participation; (3) the possible clinical findings that may result from the research and whether they will be disclosed to subjects; (4) the possibility for commercial gain from the research endeavor (as addressed in Opinion 2.08 of the Code of Medical Ethics⁹); (5) the possible conflict of interests that investigators face (as discussed in Opinion 8.0315⁹); and (6) the right to withdraw from the research at any time.

In addition, disclosure should specify what measures will be used to protect privacy. It also should include information regarding whether samples will be stored, and if so, identify (1) the scope of any additional research or intended uses foreseeable at the time the sample is collected and (2) the time and manner in which archived information and samples will be discarded.

With regard to privacy, subjects should be informed of whether the study protocol specifies what personal information will remain linked to their materials. Indeed, for follow-up purposes, some studies may require that materials contain personally identifiable information such as subjects' name or social security number (i.e., identifiable materials). Confidentiality of subjects' materials will be protected more completely, however, if data are stripped of all identifiers (i.e., de-identified). One shortcoming with this method is that completely de-identified information and samples may lose some of their utility, because they cannot be linked back to information re-

garding the materials' source of origin. Instead, it is customary to code personal identifiers, such that only the investigator can trace material back to specific individuals (i.e., coded samples). It is paramount that subjects understand, through the informed consent process, whether their material will remain personally identified, be coded, or be completely de-identified, or whether they have a choice among these options.

If data are to remain identified or coded, subjects should be offered the opportunity to be contacted by researchers in the future to share in findings and to consider participating in additional related or unrelated research. If they choose to be contacted in the future, subjects would have the responsibility to provide appropriate notification of any changes in contact information. Individuals should always be free to decide whether participation in any specific investigation is consistent with their moral beliefs and personal preferences and to refuse the use of their biological materials in research, without penalty. One important advantage of coding is that, even after their decision to participate in a study, subjects have the option to remove their information and samples from the database if they decide to withdraw from the investigation. De-identified samples do not offer such flexibility.

Waiver of consent

According to the American Society of Human Genetics' (ASHG) official statement on informed consent for genetic research, it is inappropriate to obtain a subject's blanket consent for the use of their archived information and samples in subsequent research if these materials contain information that can identify the individual.³

The National Bioethics Advisory Commission (NBAC) takes a different position.⁴ Assuming an analogy between the materials contained in DNA databanks and health care data recorded in medical records, its standards reflect federal regulations that allow research to proceed without the requirement of obtaining consent from subjects, where participation risks are no more than minimal.⁵ This position has drawn criticism from commentators who stress the unique nature of risks associated with genetic information. It is their belief that federal regulations, which were written mostly to anticipate physical risks, do not provide an adequate framework to protect subjects in genomic research.⁶ Moreover, new HIPAA regulations prohibit disclosure of any protected health information related to an individual's tissue samples without the individual's express permission.

In the case of research relying on previously archived data sets, investigators may encounter situations in which individual subjects are deceased and therefore unable to consent to unforeseen uses. In such instances, it has been argued that some research could proceed as long as the material is completely de-identified. Based on a strict interpretation of the principle of respect for autonomy, however, the argument also has been made that even de-identified materials—whether newly or previously collected—should not be used in ways to which subjects did not specifically consent and that individuals should be able to refuse to participate or to have their informa-

tion used in research that is contrary to their values and preferences.⁷

Presumed consent

Under the presumed consent standard that Iceland has adopted, the willingness of an individual to participate in research is assumed unless the individual takes formal measures to opt out. However, this standard can only function as an effective safeguard if concerned individuals are informed of the following: (1) the risks and benefits associated with the proposed research; (2) the fact that participation is optional (i.e., individuals who choose not to participate will not be penalized for their decision); (3) the appropriate steps to follow in order to opt out; (4) their status as subjects unless they formally opt out; and (5) the contact information for a person who can provide them with further clarification and answers to their questions.

In 1998, this model of presumed consent for participation in research was adopted in Iceland, a country with a universal health care system. A genetic company was granted a 12-year exclusive license by the Icelandic government to extract information from the Iceland Health Sector Database (HSD). A majority of Icelandic people voted in support of the creation of the database during a referendum, as a community consent process. As a result, competent Icelandic people were presumed willing to have information from their medical records entered into the database, unless they take necessary measures to opt out of the HSD. Since its establishment, this model had been the source of some controversy, within and outside the country. More recently, the constitutionality of the law governing the creation of the HSD was brought into question by a decision of Iceland's supreme court, stating that the HSD Act fails to protect personal privacy adequately.⁸

In the United States, where the health care system is very different and the population far more heterogeneous, presumed consent has never been seriously considered for research. Guidelines that currently govern the practice of research in the United States do not permit the use of the standard of presumed consent as an alternative to informed consent. In fact, the general attitude toward presumed consent seems to be that it is inconsistent with the principle of autonomy, upon which ethical protections for research subjects are built. This sentiment may change if developments in genomic research suggest that highly important and unique opportunities to gain new knowledge are being missed. Pressure for change—whether toward a presumed consent model or a weaker version of informed consent—may come from the biotechnology industry if it becomes too cumbersome to conduct research that adheres to current standards of consent.

CONCLUSION

The use of DNA databanks for genomic research raises new scientific possibilities as well as new challenges. Suggestions have been made to relax the standards that govern research, in an attempt to promote the acquisition of valuable information,

although strict interpretation of informed consent may be needed to protect subjects and their communities from new forms of risks. Without proper education of potential subjects, genomic research may face severe setbacks. Therefore, it is important that physicians be prepared to discuss with their patients and/or potential subjects this powerful new research tool.

Recommendations

The following safeguards should be applied to the use of databases for the purpose of population-based genomic research:

(1) Physicians who participate as investigators in genomic research should have adequate training in genomic research and related ethical issues so as to be able to discuss these issues with patients and/or potential research subjects.

(2) If research is to be conducted within a defined subset of the general population, that is, an identifiable community, then investigators should consult with the community to design a study that will minimize harm not only for individual subjects, but also for the community. When substantial opposition to the research is expressed within the community, investigators should not conduct the study. When the community supports a proposal, investigators nevertheless should obtain individual consent in the usual manner. The same procedure should be followed whether the investigators intend to collect new samples and data or whether they wish to use previously archived data sets.

(3) When obtaining the informed consent of individuals to participate in genomic research, standard informed consent requirements apply (see Opinion 2.07⁹). In addition, the following should be adhered to: (a) Special emphasis should be placed on disclosing the specific standards of privacy contained in the study: whether personal information associated with materials will be coded (i.e., encrypted so that only the investigator can trace materials back to specific individuals) or be completely de-identified (i.e., stripped of personal identifiers). (b) If data are to be coded, subjects should be informed as to whether the option exists to be contacted in the future to share in findings or to consider participating in additional research, which may relate to the current protocol or extend to other research purposes. (c) Individuals should always be free to refuse the use of their biological materials in research, without penalty. (d) Disclosure should include information about whether investigators or subjects stand to gain financially from research findings (see Opinion 2.08⁹). Such disclosure should refer to the possible conflicts of interest of the investigators (see Opinion 8.0315⁹). (e) Subjects should be informed of when, if ever, and how archived information and samples will be discarded.

(4) To strengthen the protection of confidentiality, genomic research should only be conducted using information and samples that have been coded or de-identified and that, therefore, do not directly identify the individuals from whom they were obtained (i.e., by name or social security number). Furthermore, to protect subsets of the population from such

harms as stigmatization and discrimination, demographic information not required for the study's purposes should be coded.

ACKNOWLEDGMENTS

The Council wishes to acknowledge the contributions of Bartha Maria Knoppers, JD, Matthew McGinniss, PhD, FACMG, and Beverly Woodward, PhD, in the development of this report.

References

1. Crawshaw R, Garland MJ, Hines B, Lobitz C. Oregon health decisions: An experiment with informed community consent. *JAMA* 1985;254:3213–3216.
2. McGinnis MG. The Assent of a Nation: Genethics and Iceland. *Clin Genet* 1999;55: 234–239.
3. American Society of Human Genetics statement on informed consent for genetic research. *Am J Hum Genet* 1996;59:471–474.
4. National Bioethics Advisory Commission. Research involving human biological materials: Ethical issues and policy guidance. Rockville, Md: National Bioethics Advisory Commission; 1999. Research Involving Human Biological Materials: Ethical Issues and Policy Guidance, Vol. I: Report and Recommendations of the National Bioethics Advisory Commission. Available at: <http://www.georgetown.edu/research/nrcbl/nbac/pubs.html>.
5. Common rule for the protection of human subjects. 1991. US Code of Federal Regulations. Title 45, pt. 46, sec. 110. Available at: <http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm#46.110>.
6. Greely HT. Breaking the stalemate: A prospective regulatory framework for unforeseen research uses of human tissue samples and health information. *Wake Forest Law Review* 1999;34:737–766.
7. Sade RM. Research on stored biological samples is still research. *Arch Int Med* 2002; 162:1439–1440.
8. Abbott A. Icelandic database shelved as court judges privacy in peril. *Nature* 2004; 429:118.
9. Council on Ethical and Judicial Affairs, American Medical Association. *Code of Medical Ethics, Current Opinions with Annotations*. Chicago, IL: AMA Press; 2004.