

Reproduced with permission from BNA's Health Law Reporter, 20 HLR 1570, 10/20/2011. Copyright © 2011 by The Bureau of National Affairs, Inc. (800-372-1033) <http://www.bna.com>

## Genetic Privacy in Clinical Research: The Risks to Institutions



By JILL E. ANDERSON

**R**ecent advances in genetic science have made it more crucial than ever for institutions<sup>1</sup> to address the risk of breaches in privacy when conducting clinical research. Genetic information about an individual participating in a clinical trial (a study subject) can reveal private and sometimes critical information about the study subject's past and future health. In addition, the unintended release or disclosure of private information can place a study subject at risk for genetic discrimination and certain groups at risk for stigmatization.<sup>2</sup>

Rapid advances in genetic technology have led to a significant increase in how genetic information is used. Genetic research could provide tremendous knowledge

<sup>1</sup> For purposes of this article, "institution" refers to an academic medical center, research institution, or teaching hospital which conducts clinical research.

<sup>2</sup> OFFICE OF BIOREPOSITORIES AND BIOSPECIMEN RESEARCH, NATIONAL CANCER INSTITUTE, NATIONAL INSTITUTES OF HEALTH, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES, NATIONAL CANCER INSTITUTE BEST PRACTICES FOR BIOSPECIMEN RESOURCES (June 2007), [http://biospecimens.cancer.gov/global/pdfs/NCI\\_Best\\_Practices\\_060507.pdf](http://biospecimens.cancer.gov/global/pdfs/NCI_Best_Practices_060507.pdf) [hereinafter BIOSPECIMEN BEST PRACTICES].

*Jill E. Anderson is counsel to Moses & Singer LLP's health care practice, New York City, where she counsels a variety of entities in the health care industry, including sponsors, on issues relevant to clinical research. For more information on this topic, please contact Jill at [janderson@mosessinger.com](mailto:janderson@mosessinger.com) or 212-554-7836. To learn more about Moses & Singer, please visit <http://www.mosessinger.com>.*

regarding the genetic basis of diseases and disorders, leading to earlier diagnosis and treatment of individuals with a genetic predisposition toward developing certain diseases and disorders. For example, through the identification of how an individual's gene variants influence drug metabolism, researchers may use genetic testing to predict how different individuals will respond to particular drugs.<sup>3</sup>

However, the benefits of genetic technology come with risks to breaches of the study subjects' privacy and institutions have increased liability as a result of these risks. There are several steps an institution should take to minimize its liability when conducting clinical research that involves genetic information.

### Uses of Genetic Information

There now are two federal regulations that could potentially apply to research involving genetic research: the Common Rule,<sup>4</sup> which regulates all federally funded research and sets forth the federal policy for the protection of human research subjects, and the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule, which restricts certain unauthorized uses and disclosures of patients' identifiable protected health information by covered entities. The Common Rule does not apply to a genetic study of existing data or tissue that does not involve identifiable private information<sup>5</sup> because the study is not considered to constitute human subjects research.<sup>6</sup> As such, once the determination is made that the genetic study does not involve human subjects research, no IRB oversight or informed consent is mandated. However, there still is the issue of

<sup>3</sup> Sonia Y. Hunt, Pharmacogenetics, *Personalized Medicine, and Race*, 1(1) NATURE EDUCATION (2008), available at <http://www.nature.com/scitable/topicpage/pharmacogenetics-personalized-medicine-and-race-744>.

<sup>4</sup> 45 C.F.R. pt. 46.

<sup>5</sup> 45 C.F.R. § 46.102(f) states that private information includes "information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (for example, a medical record). Private information must be individually identifiable (i.e., the identity of the subject is or may readily be ascertained by the investigator or associated with the information) in order for obtaining the information to constitute research involving human subjects."

<sup>6</sup> GUIDANCE ON RESEARCH INVOLVING CODED PRIVATE INFORMATION OR BIOLOGICAL SPECIMENS, OFFICE FOR HUMAN RESEARCH PROTECTIONS (Oct. 16, 2008), <http://www.hhs.gov/ohrp/policy/cdebiol.html> hereinafter PRIVATE INFORMATION GUIDANCE.

whether the biospecimens or information truly are not individually identifiable. For example, it is debatable whether the study subject is actually “de-identified” if his or her entire genome is sequenced. An entire genome is unique to the individual and therefore just as much an identifier as a birth date, Social Security number, or fingerprint. There are potential negative implications, including genetic discrimination and stigmatization, to the study subject if third parties have access to this genome and its associated clinical data. While the Genetic Information Nondiscrimination Act of 2008 (GINA)<sup>7</sup> removes some of the threat of genetic discrimination, the law is unclear and does not provide clear protection for a study subject about whom genetic information is derived.

GINA protects Americans from discrimination based on information derived from genetic tests. It forbids insurance companies or health plans from discriminating through reduced coverage or pricing and prohibits employers from making adverse employment decisions based on a person’s genetic predisposition to a disease. GINA amends certain laws, most notably HIPAA and the Privacy Rule promulgated under HIPAA.

Under HIPAA, the use and disclosure of certain “protected health information,” generally referred to as PHI, requires a valid consent from the research study subject. Generally, PHI means individually identifiable health information. Under GINA, the definition of “health information”<sup>8</sup> was modified to include genetic information. However, genetic information, while health information, only is covered by the Privacy Rule to the extent that it meets the definition of PHI. That is, the genetic information must be individually identifiable, maintained by a HIPAA covered entity (or business associate of a covered entity), and not otherwise fall within one of the exceptions to the definition.

“Genetic information” under GINA means information about “genetic tests” of the study subject or his/her “family member” (as defined by the act) and the manifestation of a disease or disorder in family members of such study subject. If the “genetic information” is PHI and protected by HIPAA, a valid consent must be obtained from the study subject for the use and disclosure of that “genetic information.”

The term “genetic test” means an analysis of human DNA, RNA, chromosomes, proteins, or metabolites, which detect genotypes, mutations, or chromosomal changes. An important exception to the definition of “genetic test” under GINA is that “an analysis of proteins or metabolites that is directly related to a manifested disease, disorder, or pathological condition that could reasonably be detected by a health care professional with appropriate training and expertise in the field of medicine involved” is not a genetic test.<sup>9</sup>

<sup>7</sup> Genetic Information Nondiscrimination Act of 2008, Pub. L. No. 110-233, 122 Stat. 881 (codified at 42 U.S.C. § 2000ff et seq.).

<sup>8</sup> See 45 C.F.R. § 160.103.

<sup>9</sup> Section 160.103 states that the definition of “manifestation or manifested” shall mean, with respect to a disease, disorder, or pathological condition, that an individual has been or could reasonably be diagnosed with the disease, disorder, or pathological condition by a health care professional with appropriate training and expertise in the field of medicine involved, and to further provide that a disease, disorder, or pathological condition is not manifested if the diagnosis is based principally on genetic information. *Id.*

Therefore, when a researcher runs any “genetic test” and the information from that “genetic test” can individually identify the study subject, such information would be considered PHI and would be protected under HIPAA and any use or disclosure by a covered entity would require a HIPAA authorization from the identifiable study subject. The Office of Civil Rights (OCR) has provided examples of what tests constitute a “genetic test” under the regulations.<sup>10</sup> A test to determine whether an individual has a gene variant associated with breast cancer (such as the BRCA1 or BRCA2 variant) is a genetic test. Similarly, a test to determine whether an individual has a genetic variant associated with hereditary nonpolyposis colorectal cancer is a genetic test. However, medical tests that analyze genetic material that is not of human origin, such as tests that detect the presence of viruses or bacteria in an individual, or tests that do not detect genotypes, mutations, or chromosomal changes, are not genetic tests. For example, “an HIV test, complete blood count, cholesterol test, liver function test, or test for the presence of alcohol or drugs is not a genetic test.”<sup>11</sup>

GINA has numerous ambiguities that HHS will need to address in subsequent guidance. For example, if the “genetic information” of a study subject could identify the child of a study subject, it is not clear if an authorization would be required of the child as well. Most importantly, there is no clear indication as to when “genetic information” would be deemed to be identifiable.

Current statistical techniques for analyzing genomic information have led to the possibility of determining an individual’s identity even from aggregated or pooled genetic data. For example, an individual’s genomic DNA can be detected in data containing a complex mixture of DNA from numerous individuals.<sup>12</sup> If such pooled data is published, there is the possibility that a study subject’s identity may be determined through an analysis of the published data.<sup>13</sup> As such, institutions need to clearly understand how researchers will anonymize data if the researcher intends to aggregate the data and represents the data will therefore be de-identified.

The Privacy Rule designates two ways through which a covered entity can determine that health information is de-identified. The first is the “safe harbor” approach, which permits a covered entity to consider data to be de-identified if it removes 18 types of identifiers, as listed in the Privacy Rule, and has no actual knowledge that the remaining information could be used to identify an individual, either alone or in combination with other information.<sup>14</sup> The second way is the statistical ap-

<sup>10</sup> 75 Fed. Reg. 51698, 51701 (Oct. 7, 2009), available at <http://www.hhs.gov/ocr/privacy/hipaa/understanding/special/genetic/ginanprm.pdf>.

<sup>11</sup> 26 CFR § 54.9802-3T(a)(5)(i); 29 CFR § 2590.702-1(a)(5)(i); 45 CFR § 146.122(a)(5)(i).

<sup>12</sup> Nils Homer et al., Resolving Individuals Contributing Trace Amounts of DNA to Highly Complex Mixtures Using High-Density SNP Genotyping Microarrays, 4(8) PLoS GENET E10000167 (Aug. 29, 2008), <http://www.plosgenetics.org/article/info:doi/10.1371/journal.pgen.1000167>.

<sup>13</sup> Modifications to Genome-Wide Association Studies (GWAS) Data Access, National Institutes of Health (Aug. 28, 2008), [http://grants.nih.gov/grants/gwas/data\\_sharing\\_policy\\_modifications\\_20080828.pdf](http://grants.nih.gov/grants/gwas/data_sharing_policy_modifications_20080828.pdf).

<sup>14</sup> The following identifiers of the individual or of relatives, employers, or household members of the individual must be

proach, which permits covered entities to disclose health information in any form provided that a qualified statistical or scientific expert concludes, through the use of accepted analytic techniques, that the risk the information could be used alone, or in combination with other reasonably available information, to identify the study subject is very small.<sup>15</sup>

If it is not clear that all 18 elements that could be used to identify a study subject have been removed in accordance with the Privacy Rule's de-identification standard,<sup>16</sup> institutions may need to hire a statistician to certify that there is only a very small risk that the information could be used by the recipient to identify the individual who is the subject of the information, alone or in combination with other reasonably available information.

If a study includes genetic sequencing or analysis, the informed consent form (ICF) should include information about the potential risks to the study subject posed by such research. While the frequency of harm is not known, possible consequences could include paternity determinations or potential employment or insurance discrimination if the research reveals the study subject to be at increased risk for certain diseases. In the case of genetic analyses, such risks also could pertain to the study subject's family members.

Individuals responsible for drafting the ICF and clinical trial agreement should review clinical research protocols to determine if genetic information will be collected from study subjects, whether the genetic information will enable the study subject to be individually identified, and for what purposes such information will be used. In addition, institutions should ensure the applicable informed consent form adequately describes the permitted uses and disclosures of the genetic information.

The informed consent regulations require that the ICF have a statement describing the extent to which data that identifies the study subject will be kept confi-

removed to achieve the "safe harbor" method of de-identification: (A) names; (B) all geographic subdivisions smaller than a state, including street address, city, county, precinct, ZIP code, and their equivalent geocodes, except for the initial three digits of a ZIP code if, according to the current publicly available data from the Bureau of Census (1) the geographic units formed by combining all ZIP codes with the same three initial digits contains more than 20,000 people; and (2) the initial three digits of a ZIP code for all such geographic units containing 20,000 or fewer people is changed to 000; (C) all elements of dates (except year) for dates directly related to the individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older; (D) telephone numbers; (E) fax numbers; (F) electronic mail addresses; (G) Social Security numbers; (H) medical record numbers; (I) health plan beneficiary numbers; (J) account numbers; (K) certificate/license numbers; (L) vehicle identifiers and serial numbers, including license plate numbers; (M) device identifiers and serial numbers; (N) web universal resource locators (URLs); (O) internet protocol (IP) address numbers; (P) biometric identifiers, including finger and voice prints; and (Q) full face photographic images and any comparable images; and any other unique identifying number, characteristic, or code, except as permitted for re-identification purposes provided certain conditions are met.

<sup>15</sup> 45 C.F.R. § 164.514(b).

<sup>16</sup> See § 164.514 under the Privacy Rule.

dential.<sup>17</sup> For genetic studies, the National Institutes of Health (NIH) recommends language on the National Human Genome Research Institute (NHGRI) website.

Genetic research can have many goals, including DNA sequencing, genotyping or studying genetic variations across populations. Based on the type of genetic research being conducted, institutions will need to modify their consent language to address concerns specific to such research. As part of the NHGRI, the NIH has published many different informed consent models for use in genetic research. Institutions may wish to use one of these models to develop their own template. The model consents can be located by accessing the NHGRI website at <http://www.genome.gov>.

### *Data That Are Linked With Biospecimens*

While the Privacy Rule<sup>18</sup> does not specifically state that it applies to biospecimens directly, a study subject's tissue samples could contain information which researchers may be able to use to re-identify the study subject. Federal regulations require informed consent, a valid HIPAA authorization and IRB review for research involving human subjects.<sup>19</sup> Excepted from these requirements is research involving "the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects."<sup>20</sup> In addition, PHI under the Privacy Rule does not include de-identified information<sup>21</sup> or biospecimens with no accompanying information, such as an accession number or code number that may be linked to an identifier. Therefore, if a researcher receives only de-identified biospecimens or study subject information, no consent or HIPAA authorization is required for research conducted on these samples or study subject information, although some states still require IRB review. The collection of longitudinal data with the associated biospecimens may cause the data to be more readily identifiable. Longitudinal data<sup>22</sup> may include age, ethnic background, diagnosis, disease history, medical treatments, and response to treatments. These data may, in combination, allow the re-identification of the study subject.<sup>23</sup>

Biorepositories collect, store, process, and distribute biospecimens and associated study subject data to allow researchers to study, for example, the connection between molecular information and study subjects' clinical responses. Federal law governing this area is

<sup>17</sup> 45 C.F.R. § 46.116(a)(5).

<sup>18</sup> 45 C.F.R. pt. 160; 45 C.F.R. § 164.102 et seq.; 45 C.F.R. § 164.500 et seq.

<sup>19</sup> 45 CFR pt. 46; 21 C.F.R. pts. 50, 56; 45 C.F.R. pt. 160; 45 C.F.R. § 164.102 et seq.; 45 C.F.R. § 164.500 et seq.

<sup>20</sup> 45 C.F.R. § 46.101(b)(4).

<sup>21</sup> 45 C.F.R. §§ 164.502(d)(2), 164.514(a) and (b).

<sup>22</sup> "Longitudinal data" comprise outcome measures or responses from the same subjects across multiple time points along with study subject characteristic measures. The data may be used to answer questions regarding the change or progression of the disease over time and/or potential risk factors for the disease. The existing data could be used to identify subgroups with different background characteristics or traits. Li, H.-I., *Longitudinal Data*, WILEY ENCYCLOPEDIA OF CLINICAL TRIALS (2008).

<sup>23</sup> BIOSPECIMEN BEST PRACTICES.

ambiguous and often confusing to researchers because it raises several issues related to ownership of specimens, informed consent, genetic privacy, and retention of DNA specimens and associated information.

The Office of Human Research Protections (OHRP) has issued guidance on regulatory requirements that must be satisfied by biorepositories.<sup>24</sup>

As part of the guidance, OHRP recommends that, included among the basic elements of informed consent, there should be a clear description of (i) the operation of the biorepository; (ii) the specific types of research to be conducted; (iii) conditions under which data and specimens will be released to recipient-investigators; and (iv) procedures for protecting the privacy of study subjects and maintaining the confidentiality of data. Informed consent information describing the nature and purposes of the research should be as specific as possible. In addition, when future genetic research is anticipated, informed consent information should include information about the consequences of DNA profiling (e.g., regarding possible paternity determinations).

Under the definition of human subject, obtaining<sup>25</sup> identifiable private information or identifiable specimens for research purposes constitutes human subjects research.<sup>26</sup> The OHRP has stated that, under certain limited conditions, research involving only coded private information or specimens is not human subjects research.<sup>27</sup> In addition, the Privacy Rule<sup>28</sup> permits covered entities under the rule to determine that health information is de-identified even if the health information has been assigned, and retains, a code or other means of record identification, provided that:

- (i) the code is not derived from or related to the information about the individual;
- (ii) the code could not be translated to identify the individual; and
- (iii) the covered entity under the Privacy Rule does not use or disclose the code for other purposes or disclose the mechanism for re-identification.<sup>29</sup>

Regarding condition (i) above, in contrast to the Privacy Rule, information that is linked with a code derived from identifying information or related to information about the study subject is not considered to be individually identifiable under the Common Rule, if the

researcher cannot readily ascertain the identity of the individual(s) to whom the coded private information or biospecimen pertains. Therefore, some coded information, in which the code has been derived from identifying information linked to or related to the individual, would be individually identifiable under the Privacy Rule, but might not be individually identifiable under the Common Rule.

Institutions which conduct research on biospecimens need to develop clear policies and procedures governing the privacy and security of biospecimens and associated clinical data. If the biospecimens or associated data are coded, the institution storing such biospecimens or data should have a policy governing the maintenance and use of the code. Data associated with biospecimens should be coded and only the researcher should have access to a secure link to the code that identifies the study subject. If an institution or researcher wishes to transfer the biospecimens to third parties, it should execute an agreement with the third party and include language requiring the third party to only use such biospecimens for specific purposes, as permitted by the informed consent and authorization taken from the study subject.

Although all research uses of biospecimens may not be known at the time the biospecimen is collected, the informed consent should describe the nature and purposes of the research as specifically as possible.<sup>30</sup> Researchers, including those at biotechnology or pharmaceutical companies, who use the biorepository should sign a data use agreement agreeing to not to attempt to re-identify patients. In addition, the consent form should contain an “opt-out” provision allowing patients to opt-out of having their biospecimens and/or data placed in the biorepository.

Specific actions that an institution or recipient researcher, as appropriate, could take if a study subject discontinues participation are: (i) cease using the study subject’s individually identifiable specimens and private information in the study; (ii) remove the study subject’s individually identifiable information from the biospecimens and eliminate private information (if doing so would render the specimen not individually identifiable to the investigators); or (iii) destroy the study subject’s individually identifiable specimen and private information.<sup>31</sup>

Institutions may take additional steps to ensure that the information is properly de-identified. For example, institutions could develop algorithms for procedures for de-identification. To avoid potential liability for third party use of biorepository samples which were taken from study subjects as part of a research study, institutions should specify in the clinical research agreement that the sponsor only will use the data in accordance with the informed consent form and authorization, or as required by law. In addition, the clinical research agreement should require the sponsor to be responsible for the sponsor’s disclosure of the study subject’s information to any third party. Finally, institutions should ensure the informed consent and authorization clearly disclose how the study subject’s genetic information may be used and by whom.

Institutions can mitigate the risks of a privacy breach by knowing potential uses of genetic information, ad-

<sup>24</sup> See OHRP guidance entitled, “Issues to Consider in the Research Use of Stored Data or Tissues” <http://www.hhs.gov/ohrp/policy/reposit.html> and “Research Involving Coded Private Information or Biological Specimens” <http://www.hhs.gov/ohrp/policy/cdebiol.html>.

<sup>25</sup> “Obtaining” identifiable private information or identifiable specimens includes, but is not limited to:

(1) using, studying, or analyzing for research purposes identifiable private information or identifiable specimens that have been provided to investigators from any source; and

(2) using, studying, or analyzing for research purposes identifiable private information or identifiable specimens that already were in the possession of the investigator.

<sup>26</sup> 45 C.F.R. § 46.102(f).

<sup>27</sup> Office for Human Research Protections, U.S. Department of Health and Human Services (HHS), “Guidance on research involving coded private information or biological specimens” (HHS, Washington, D.C., Oct. 16, 2008).

<sup>28</sup> See 45 C.F.R. pt. 160; 45 C.F.R. § 164.102 et seq.; 45 C.F.R. § 164.500 et seq.

<sup>29</sup> See INSTITUTIONAL REVIEW BOARDS AND THE HIPAA PRIVACY RULE GUIDANCE, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES 6 (August 2003), [http://privacyruleandresearch.nih.gov/pdf/IRB\\_Factsheet.pdf](http://privacyruleandresearch.nih.gov/pdf/IRB_Factsheet.pdf).

<sup>30</sup> *Id.*

<sup>31</sup> *Id.*

describing how those uses could identify a study subject and ensuring the risks of identification are appropriately disclosed in the informed consent form.

# MOSES & SINGER LLP

---

## Disclaimer

Viewing this or contacting Moses & Singer LLP does not create an attorney-client relationship.

This is intended as a general comment on certain developments in the law. It does not contain a complete legal analysis or constitute an opinion of Moses & Singer LLP or any member of the firm on the legal issues herein described. This contains information that may be modified or rendered incorrect by future legislative or judicial developments. It is recommended that readers not rely on this general guide in structuring or analyzing individual transactions or matters but that professional advice be sought in connection with any such transaction or matter.

## Attorney Advertising

It is possible that under the laws, rules or regulations of certain jurisdictions, this may be construed as an advertisement or solicitation.

Copyright © 2011 Moses & Singer LLP  
All Rights Reserved