Comments of the World Privacy Forum

To The Department of Health and Human Services, Office of the Secretary, and Food and Drug Administration

Regarding the Advance notice of proposed rulemaking: Human Subjects Research Protections: Enhancing Protections for Research Subjects and Reducing Burden, Delay, and Ambiguity for Investigators , Docket ID number HHS–OPHS–2011–0005

Via Regulations.gov

October 18, 2011

> **Re: Advance notice of proposed rulemaking: Human Subjects Research Protections: Enhancing Protections for Research Subjects and Reducing Burden, Delay, and Ambiguity for Investigators , Docket ID number HHS–OPHS–2011–0005**

To the Office of the Secretary, US Department of Health and Human Services:

The World Privacy Forum appreciates the opportunity to comment on the Advance Notice of Proposed Rulemaking (ANPRM) on human subjects research protections that appeared in 76 Federal Register 44512 (July 26, 2011). These comments mostly address privacy issues raised by the ANPRM. The World Privacy Forum is a 501(c)(3) non-partisan public interest research group based in California. Our funding is from foundation grants, cy pres awards, and individual donations. We focus on conducting in-depth research on emerging and contemporary privacy issues as well as on consumer education. A core area of our work is in health care privacy issues, among other topics.[1]

[1] The World Privacy Forum website contains more information about our organization, our publications, and activities. See http://www.worldprivacyforum.org.

I. Information Risks in General

The ANPRM focuses appropriately on the informational risks faced by research subjects. On page 44516, the notice states:

> Informational risks derive from inappropriate use or disclosure of information, which could be harmful to the study subjects or groups. For instance, disclosure of illegal behavior, substance abuse, or chronic illness might jeopardize current or future employment, or cause emotional or social harm. In general, informational risks are correlated with the nature of the information and the degree of identifiability of the information. The majority of unauthorized disclosures of identifiable health information from investigators occur due to inadequate data security.

This view is too narrow. While avoiding informational risks is a proper objective, the better and broader goal should be providing **Fair Information Practices** for data subjects. This objective reflects the increasingly common perspective that privacy protections should meet the standards of Fair Information Practices (FIPs).[2] Indeed, FIPs had their origins in the Department, and they form the basis for the Privacy Act of 1974 and for the HIPAA Privacy Rule.[3]

It is important to take a proper view of the privacy interests of data subjects from the start. Informational risks are not limited to inappropriate use or disclosure of information. Focusing only on **use and disclosure** ignores the other elements that are just as essential to addressing the privacy interests of data subjects. Admittedly, the ANPRM and the existing rule address some of these other interests, but not all, and not clearly.

Starting with an incomplete concept of privacy leads to an incomplete approach. For example, the basic compilation of personal information by a record keeper (who may or may not be acting with the knowledge and consent of a data subject) exposes the data subject to an informational risk *regardless of the use and disclosure rules that apply*. These other risks are not at zero just because there are limits on use and disclosure. Some disclosures are beyond the control of any record keeper. Any litigant or agency with a subpoena and any law enforcement officer with a search warrant may be able to force disclosure of a record, regardless of the record keeper's disclosure policy. Had the records not been compiled in the first place, these risks might not exist.

---

[2] The Obama Administration appears to be broadly supportive of FIPs as a basis for privacy policy. See, e.g., the Department of Commerce Internet Task Force, *Commercial Data Privacy and Innovation in the Internet Economy: A Dynamic Policy Framework* (2011), at http://www.commerce.gov/sites/default/files/documents/2010/december/iptf-privacy-green-paper.pdf.

[3] Department of Health and Human Services, Final Rule, Standards for Privacy of Individually Identifiable Health Information, 65 Federal Register 82462, 82464 (Dec. 28, 2000) at http://www.gpo.gov/fdsys/pkg/FR-2000-12-28/pdf/00-32678.pdf ("This final rule establishes, for the first time, a set of basic national privacy standards and fair information practices that provides all Americans with a basic level of protection and peace of mind that is essential to their full participation in their care."). For a history of FIPs, see Robert Gellman, *Fair Information Practices: A Basic History,* at http://bobgellman.com/rg-files/rg-resources.html. It is noteworthy that Congress directed the Department of Homeland Security to assure compliance with FIPs. 6 U.S.C. §142.

Other risks of improper use and disclosure remain – again, regardless of the rules that are supposed to apply. Security breaches occur even when good privacy and security policies are in place.[4] Disclosure of an individual's medical history to a neighbor who happens to be working on a research project is a risk that disclosure policies cannot cure. Insufficient attention to data quality can affect a data subject in a variety of ways. A data subject also has an interest in (and in many circumstances a right to) access and correction of personal information, limits on collection of personal information, and other FIPs elements. How these elements may apply in a research context is open to discussion, but all elements of FIPs need to be addressed in some appropriate way. The point is that privacy and informational risks are not just a matter of *use* and *disclosure*.

There is a broader point to be made here as well. Privacy is not just about avoiding harm. Some would like to approach privacy by arguing that if there is no demonstrable harm (often financial harm), then there is no privacy interest at all. Those who profit commercially from the exploitation of personal information are happy with a harms based approach because they know that individuals have difficulty proving harm to the satisfaction of a court of law. The better approach is to view privacy as a right of individuals. Europeans treat privacy as a fundamental human right, and they regulate record keepers to protect privacy accordingly. HIPAA treats privacy as a right, and the HIPAA privacy rule establishes standards that protect individuals regardless of evidence of actual harm. Human subjects protection needs to start from a premise that privacy is a right of individuals that requires suitable and full protection.

II. Responses to Specific ANPRM Questions

In the responses below, we have replicated the relevant ANPRM question and have responded directly to that question. In this comment, we address the ANPRM questions affecting privacy. We have used the ANPRM question numbers in this document.

A. Minimal Risk

Question 1: Is the current definition of ''minimal risk'' in the regulations (45 CFR 46.102(i) – research activities where ''the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests'') – appropriate? If not, how should it be changed?

Response: It is not clear how or if the Common Rule's concept of minimal risk applies or should apply to informational risks. The risks to privacy *ordinarily encountered in daily life* are significant, poorly understood, and often hidden from view. Current debates over the use of information about Internet or cell phone users illustrate the point. Regular news reports highlight novel and often-secret data collection methods used by commercial companies in the pursuit of

[4] See, e.g., *Chronology of Data Breaches Security Breaches 2005 - Present* maintained by the Privacy Rights Clearinghouse at https://www.privacyrights.org/data-breach.

consumer profiles, more effective advertising, or increased profits.[5] Studies show broad gaps in popular knowledge about privacy.[6] Current debates in Congress and state legislatures provide additional evidence of the need for better privacy rules.

Whether the privacy risks *ordinarily encountered in daily life* is a fair standard for measuring information risks in research is doubtful. Researchers should be held to higher standards than apply today in the unregulated marketplace for consumer data.

Ongoing debates over the proper scope of the Fourth Amendment provide some lessons. A commonly used test to assess when privacy should fall under the protections of the Fourth Amendment comes from a concurring opinion in a seminal Supreme Court decision.[7] Under the test, a *reasonable expectation of privacy* exists if (1) a person has exhibited an actual (subjective) expectation of privacy; and (2) that expectation is one that society is prepared to recognize as reasonable. When this test is satisfied, a government search or surveillance activity that violates the reasonable expectation of privacy falls under the Fourth Amendment. The scope and applicability of the Fourth Amendment is not an immediate concern in the human subjects context.

The Fourth Amendment privacy test based on expectations is not much different from the standard in the Common Rule because both rely upon what is ordinarily encountered, routine, or expected. This brings us to the point. A well-recognized problem with the reasonable expectation of privacy test is the "silent ability of technology to erode our expectations of privacy."[8] That is precisely the problem if the standard for minimal risk in the existing rule applies to informational risks. What is happening on the Internet and elsewhere is changing and eroding expectations of privacy. The judicial *expectation of privacy* test is now widely criticized for its lack of any real objective or fixed standard. No matter how the Fourth Amendment is interpreted by the courts, we need to do better in the research context. We cannot allow the *anything goes* practice of commercial Internet advertisers to affect standard for research conduct.

---

[5] See, e.g., Center for Digital Democracy, Examination of Online Direct-to-Consumer Prescription Drug *Promotion* (2011) (Comments to Food and Drug Administration) ("Ongoing tracking and targeting of health users across the digital marketing system is a major concern not only for patient privacy, but also because such data is used to better hone campaigns designed to influence consumers in a myriad of ways. Health marketers strive to harness the data collection and analysis capabilities of online advertising in order to foster greater demand for prescription drugs. Unfortunately, little information is currently provided on what is being collected from users of health-related sites and how such data are used."), at http://www.centerfordigitaldemocracy.org/sites/default/files/FDAComments062711final.pdf.

[6] See Chris Jay Hoofnagle et al, *How Different are Young Adults from Older Adults When it Comes to Information Privacy Attitudes and Policies?* (2010) ("42 percent of young Americans answered all of our five online privacy questions incorrectly. 88 percent answered only two or fewer correctly. The problem is even more pronounced when presented with offline privacy issues – post hoc analysis showed that young Americans were more likely to answer no questions correctly than any other age group."), at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1589864. In other words, young and old alike did poorly on the test of privacy knowledge.

[7] *Katz v. United States*, 389 U.S. 347, 361 (1967) (Harlan, J., concurring).

[8] See, e.g., Paul Schwartz, *Privacy and Participation: Personal Information and Public Sector Regulation in the United States*, 80 Iowa Law Review 553, 573 (1995).

If researchers and Institutional Research Boards, or IRBs, are left unguided to apply the *ordinarily encountered in daily life* to informational risks, they may simply conclude that the informational risks of any given research project are minimal because "everybody is on Facebook" anyway or because Google does the same thing. **The activities that search engines and other websites do to monitor users of their Internet services is not a standard that should be relevant to research privacy.** Similarly, researchers might take note of the many websites where patients publicly share their personal medical history and decide that a research project poses relatively fewer privacy challenges and therefore their projects must be okay. It is too easy for someone who wants to use another person's data to find a justification if the standard is vague and too flexible.

Informational risks in research must be measured against a firm standard, one that is not affected by every change in technology or commercial practice. For example, the HIPAA privacy standard establishes a firm set of Fair Information Practices. While there is considerable flexibility in the application of the HIPAA privacy rules in some contexts, the standards themselves are not subject to change because of external factors. Patients can expect the HIPAA standards to protect their health information in the same way, regardless of what happens on Twitter.

The same should be true for human subjects research. The need for a baseline of privacy protection must be a constant for research even though the degree of informational risk can vary from project to project. The need for rules governing collection, use, and disclosure is constant. The need for openness and accountability is a constant. The need to consider individual participation rights (access and correction) is a constant. Thus, whatever the risk involved in a given project, the need for sufficient privacy protections for personally identifiable information is a constant.

In general, the ANPRM's suggestion that standardized data protections may be more effective than IRB review at minimizing informational risks is a fair conclusion in some and perhaps most cases. The research community in general has not excelled in addressing privacy. Many research projects do not typically have written privacy or security policies.[9] Whether this is the fault of the IRBs or the researchers themselves is not important. The current system does not provide the protections that research subjects deserve and something else is required. Informational risks may be the hardest for research subjects to understand and evaluate. The same may be said for IRBs and of researchers themselves.

The need is not to determine if there is minimal informational risk for any given project. The need is to make sure that the privacy and security protections for a project reasonably address the risks involved. The Department must resist any suggestion that a determination of minimal informational risk means that privacy and security can be ignored. It is inappropriate to apply the notion of minimal risk as a threshold measure of informational risk. *When there is Personally*

[9] Where institutions subject to HIPAA conduct research, the situation may be better because HIPAA privacy and security rules may apply directly and because employees with the requisite experience may be available to help researchers comply. For many other types of research conducted in other organizations lacking in external privacy and security standards, however, privacy and security are not likely to receive adequate attention or documentation.

Identifiable Information (PII), there is informational risk that must be addressed in some fashion.

Researchers should be required to apply standard data protection policies and practices as appropriate for their activities. To simplify and standardize the task, the Department should consider commissioning a **formbook of privacy policies** for use by a wide range of researchers. Each of the FIPs would be represented in the formbook with multiple implementations.[10] Researchers could then select from already prepared policies without the need to reinvent the privacy wheel each time. The creation of a privacy policy for many research projects could be as simple as ordering a meal from a menu. For security, the HIPAA rule already provides a standard that can be applied in much the same way.

To summarize, the issue is not whether there is minimal informational risk. When there is PII, there is informational risk. The privacy and security policies that apply to any given project may properly vary depending on the specifics, but there must be privacy and security policies that address the risks involved. Minimal risk may not be a meaningful threshold test for informational risk. No finding of *minimal risk* based standards found in ordinary life should exempt any research project with PII or potentially identifiable data from formally considering and addressing informational risks.

B. Survey Instruments and Minimal Risk

Question 6: Are there survey instruments or specific types of questions that should be classified as greater than minimal risk? How should the characteristics of the study population (e.g. mental health patients) be taken into consideration in the risk assessment?

Response: The ANPRM suggests a number of changes to the existing rule about exempt categories of research, including allowing exemptions for research that could reasonably place research subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation. The stated reason is these *significant risk* studies would be conducted with competent adults and because these studies would be subject to standard data security and information protection standards.

The notion that competent adults participating in research will understand the consequences of participation – especially for research that may result in the *significant risks* identified in the current rule – is false. One only has to look at the enormous sums lost by "competent adults" to electronic mail scams offering lottery winnings, bequests, and the like. Similarly, few users of social networks reasonably understand the consequences of participation or the privacy policies that apply. Do the competent adults reading these comments who also use Facebook really understand the privacy settings available to them, know how to use the settings, and truly

[10] For example, researchers involved in international activities may need to adopt privacy policies that meet the standards of the European Union. A formbook should offer an EU-compliant version. There could even be a need to have seek approval from the EU of a research "industry" standard that meets EU requirements for those researchers who need to comply. Any type of privacy standard or documentation that can be produced once and applied many times will be highly cost-effective.

appreciate the consequences of not using the settings? Even if you do, would you make the same assessment about your brother-in-law or nephew?

It is well known that many people do not always read or understand consent forms, and most have a dim understanding of legal protections and the lack thereof. Look at the losses suffered by "competent adults" (some qualified professionals) entering into mortgages, issuing mortgages, and buying mortgage paper. For all the "competence" present, there were tens of billions of dollars of losses and many ruined financial lives.

For a data subject who did not truly understand or think through the consequences of participation in a research study, standard data security and information protections will do little or nothing to alleviate liability, protect financial standing, preserve employability, or defend reputation. The data subject who casually agreed to participate in a research project and then a year later is sued for divorce, faces a child custody contest, or is arrested may never have envisioned that the research record would be used in that subsequent proceeding.

Frankly, review of a research project by an IRB may not protect individuals or enhance their understanding either. It is far from clear that the average IRB member, principal investigator, or HHS employee would do well in a basic quiz on the legal protections and lack thereof for data held in a research project. However, a review by an IRB is better than relying solely on the knowledge and experience of the average "competent adult", who may be an individual whose age, experience, overall reading ability, fluency in English, and attention span may be anywhere within a wide range. One way to improve IRB review of projects with privacy consequences is to require that each IRB include a privacy expert as one of its members. Privacy expertise is much easier to find today, if for no other reason than because of HIPAA.

Requiring security and privacy standards for all projects that qualify for exemption will still help some if the privacy protections offered are robust and are honored by the researcher. However, especially for projects that could reasonably place a data subject at risk of criminal or civil liability or be damaging to the subject's financial standing, employability, or reputation, additional protections are needed. Unfortunately, HHS can only do so much in its rules to keep research data from third parties. As noted in a previous answer, anyone with a subpoena or search warrant may obtain data from any record keeper.

Not all research that involves *significant risk* subjects will necessarily require IRB review. For example, if a research activity (e.g., a focus group) about use of illegal drugs does not record or maintain any identifiable or potentially identifiable information about a research subject, the concerns about subsequent use of information against the research subjects may not be present. There may be other reasonable exceptions as well.

There is another aspect of *significant risk*s research that requires additional consideration and review. Who is the investigator? The experience and credentials of the investigator should make a difference. It is unreasonable to treat all *significant risk* projects the same whether they are conducted by a Ph.D researcher with years of experience, by a first year graduate student, or by a college sophomore. The researcher may or may not be fully qualified.

For all of these reasons, some projects involving competent adults require some independent oversight by an IRB.

Regardless how HHS approaches that aspect of this issue, there is more to be done. One thing that HHS can do is to **require use of certificate of confidentiality programs**[11] to provide a degree of additional protection. The ANPRM recognizes some of the limits of certificate programs (p. 44525), but certificates of confidentiality offer some measure of protection, however imperfect.

If HHS allows an exemption for a research project that collects and maintains information that could significantly damage research participants, HHS should also require the project to obtain a certificate of confidentiality. Requiring a certificate would take an additional step in the direction of protecting research subjects. A certificate would result in some protection for data subjects.

HHS can and should go further.

- First, HHS should require every research project subject to the Common Rule to obtain a certificate of confidentiality if the project collects and maintains any PII. The administrative burden for HHS could be met by designing a website that would automatically issue a certificate to any project that meets basic qualifications.

- Second, anyone holding a certificate should be required to notify a research subject if the subject's record is sought through legal process. Some exceptions to notice may be appropriate.

- Third, anyone holding a certificate should be required to refuse to disclose records protected by a certificate.

- Fourth, institutions sponsoring research should be required to bear the burden of challenging subpoenas in court. HHS may be able to accomplish some of these goals through regulation or through its own grant instruments. If not, then HHS should propose legislation to expand the protections provided by certificates and to clarify the obligations of principal investigators and their institutions.

HHS can take an important additional step regarding certificates of confidentiality. It could through the Common Rule or otherwise require researchers with a certificate of confidentiality to report to HHS or otherwise make public any disclosure request that they receive *before responding to the request*. Disclosure would be required both for compelled requests (e.g., subpoena) and non-compelled requests. Disclosure of both categories of requests would be valuable. For compelled requests, public reporting might well deter some requesters, would allow researchers to find assistance with legal and other responses, and make it more likely that researchers will rely upon the legal defense that they have a result of the certificate.

[11] See, e.g., 42 U.S.C. § 241(d).

For non-compelled requests, public reporting of the requests is even more essential. A researcher with a certificate of confidentiality may be tempted to avoid the expense of resisting a subpoena by making a voluntary disclosure. If public reporting of any proposed voluntary disclosure were required, a researcher would likely find it much more difficult to evade responsibility for protecting the privacy of research subjects.[12] Public reporting might also inhibit those seeking records from making requests that would only call attention to the abuse of research records that their requests entailed. Public reporting by researchers could be enforced by making failure to report a factor in making decisions about future grants.

While certificates offer important protections, they will only do so much to protect data subjects. They will do nothing to protect a "competent adult" who casually chooses to participate in a research project by an unscrupulous researcher or poorly supervised student that exposes participants to undue risks for a research project will few or no countervailing benefits. There is a need for third party review to protect adults, no matter how competent they may appear, from themselves. Competent adults often seek advice from others when engaging in activities that may be hazardous to their personal or financial well being. Participation in "risky" research is no different.

A separate area of concern involves research projects that include any possibility of recontact of the research subject. Recontact raises a host of complex ethical issues beyond the scope of this response. **However, a research project that involves recontact requires more careful review by an IRB regardless of participation by a "competent adult" and regardless of the presence of a privacy and security policy.** Few individuals are likely to understand the consequences of recontact without a greater explanation, and IRBs can prohibit or impose fair limitations on recontact that will protect the interests of individuals who are not capable of foreseeing the risks or negotiating their own terms for recontact. The growing number of electronic health records and networks is likely to make recontact easier to undertake, and researchers may be more likely to seek recontact (often for perfectly reasonably purposes) because the barriers are reduced.

C. Consent and use of data collected for non-research purposes

Question 45: Under what circumstances should future research use of data initially collected for non-research purposes require informed consent? Should consent requirements vary based on the likelihood of identifying a research subject? Are there other circumstances in which it should not be necessary to obtain additional consent for the research use of currently available data that were collected for a purpose other than the currently proposed research?

Response: Here and in many other places, the ANPRM assumes that all research is of equal and high value. This is clearly not the case. An experienced IRB member or researcher can readily sort research proposals by their prospective value based on a number of factors.

[12] Public reporting should probably be limited to demands or requests for records about multiple individuals. Publicly disclosing a request for the records of a particular individual could negatively affect that individual's privacy interest, and personal notice to the individual should be sufficient . Public reporting may also be unnecessary in some cases when a researcher refuses a request.

Research is conducted for a wide variety of purposes by researchers with highly variable skills. All research is subject to the standard barriers to completion or success, including bad research design; loss of interest by the researchers themselves; loss of funding; loss of data; failure to publish results; transfer, retirement, or death of personnel; and the like. The ANPRM quotes (p. 44524) an IOM report that says it is important to "facilitate *important* health research by maximizing the usefulness" (emphasis added). Accepting that statement as fair, it follows that facilitating *unimportant* health research deserves a lower priority, if any priority at all.

The assumption that there should be one rule covering all research – regardless of its purpose, importance, quality, or prospects – results in a policy that overvalues research and undervalues individual interests. That result is especially likely when decisions are being made about the use of data originally collected for non-research purposes.

Striking a fair balance between all the interests involved is admittedly not simple. For a long time, we have allowed IRBs to make judgments about that balance. While IRBs have their problems, it is hard to envision a system that allows the use of personal data without consent that does not involve an independent third party making a judgment. If researchers are allowed to directly place their thumbs on the scale more than they do today, then only the interests of researchers will matter, and the privacy and other interests of data subjects will be heavily discounted at best and ignored entirely at worse.

The risks here are not just to data subjects. It will only take a single, high-profile case to threaten the entire research enterprise. Consider the individual harmed in some way because of research use of his or her data without the individual's knowledge or consent. Perhaps that individual cannot find a job, loses health insurance, faces a ruined reputation, finds his relatives condemned to any of these outcomes, or otherwise. Some may be personally offended by the use of their records or specimens for research that they object to on moral, ethical, religious, or political grounds.

If one aggrieved individual manages to attract sufficient press attention to the essential unfairness of the activity, the result could broadly change the way that research is conducted in the United States. Imagine that individual with a compelling story doing a round of talk shows warning people about participating in research.

Lawsuits are also foreseeable. So is legislation. A single news story could result in restrictive legislation that mandates affirmative consent for research use of data or sample in most or all cases. A single lawsuit against a researcher, research sponsor, or data provider could freeze research activities for years, as lawyers advise institutions against doing anything that might result in liability until the lawsuit is resolved. These are the risks faced by researchers and research sponsors in general.

The American public already generally believes that individual consent should be required before their data can be used for research. Anyone can massage polls and ask loaded questions to obtain a different answer, but public views have been quite consistent over the years. The use of

biospecimens for research is not likely to change existing public sentiments. It may exacerbate them.

Requiring researchers to have privacy and security policies is not enough because the policies will not provide sufficient protections for individuals whose data and lives are at risk. Mandating privacy and security policies – and for present purposes, we can temporarily overlook the fact that the policies that will assuredly be followed indifferently by many researchers – helps some but will not guarantee protections for individuals. Requiring researchers to adopt policies is necessary, but it is not sufficient.

Any process that allows a researcher to use PII that was not collected with adequate consent from the data subject must, at a minimum, be able to guarantee that the data subject will not be placed at greater legal or other risk because of the researcher's activities. The physician's *do no harm* principle is appropriate here. The best we can do today is to use the existing certificate of confidentiality program that provides researchers with the tools to protect data against many unrelated third party demands for data. Any researcher using data initially collected for non-research purposes should be required to obtain a certificate of confidentiality. Certificates provide some measure of protection for data subjects whose information and biospecimens are used without their knowledge or consent. The response earlier in our comments to ANPRM's question No. 6 offers additional suggestions about the use and terms of certificates of confidentiality to address their shortcomings. In addition, researchers and data sources may have to confront the possibility that their actions that result in harm to a data subject – especially when data is used and shared without knowledge or consent of the data subject – may expose research supporter to liability.

There is much else to debate about when and how researchers should be allowed access to data collected for non-research purposes. Some data will be collected under specific terms that may preclude research use or allow some activities under conditions that differ widely. It is difficult to treat all *data initially collected for non-research purposes* as one category. Some data will be subject to state or federal legislation, some will be subject to ethical limitations, some will be subject to privacy policies, some to common understandings about secondary use. The suggestion here for using certificates of confidentiality may help data subjects, but it does nothing to address the multiple types of data and circumstances that apply to *data initially collected for non-research purposes*. The problem needs more nuanced thinking.

D. Consent and analysis of data collected for different research purposes

Question 46: Under what circumstances should unanticipated future analysis of data that were collected for a different research purpose be permitted without consent? Should consent requirements vary based on the likelihood of identifying a research subject?

Response: This is a difficult question to answer in the abstract. Much depends on what the data subject was told at the time of collection. There are some simple rules that must be followed: any promise made or assurance given must be strictly followed. If the data subject was told that *your data will never be shared*, then the data may never be shared, whether it is identifiable or not. If a different promise was made, then the specific commitment must be honored. If a data

subject was given a choice and exercised it, that choice must be respected. Identifiability is not a factor unless the promise to the data subject said that *identifiable data will never be shared*. Researchers must be held to their commitments, even when those commitments were not prepared with adequate foresight.

Part of the purpose of research privacy policies is to oblige researchers to think through and state their plans with greater precision. Too much data collection goes on without much thought about future uses. Researchers need help in thinking through these matters in advance.

Promises are complicated. In the early days of the Internet, websites often said *we will never share your data with anyone*. Promises of this type are rare now because they cannot be kept. Many record keepers share with contractors, auditors, lawyers, cloud computing services, and others in the ordinary course of operations. Data may have to be provided to anyone with a subpoena, and police with a search warrant can seize anything covered by the warrant. Today, privacy policies are more elaborate and more detailed because the world in complicated, and absolutes are few.

The research community has, for the most part, yet to face the complexities of privacy. The earlier suggestion (see response to question 1) for the preparation of a formbook of standard policies would help by creating nuanced alternatives that would make help to ensure that policies and promises are realistic. Standard approaches would also discourage adoption of privacy policy that are so open-ended that they would allow researchers to disclose any data to anyone without any standard at all. In fact, HHS should take express steps to expressly forbid abuse of this type. We already know that many people wrongly believe that the mere presence of a privacy policy at an online site means that the site cannot share data with a third party.[13] If researchers use a new requirement for a privacy policy to mislead (intentionally or otherwise) people into thinking that the policy means that data will not be shared with third parties, it is possible that the requirement itself could make things worse overall. That result must be avoided. There must be some absolute limits. For example, no researcher should be allowed to share PII with marketers without knowing, express, written, time-limited consent from a data subject.

What should happen with data originally collected without data subject knowledge or consent? In many ways, the same answer just given applies here. Research activities involving PII should have privacy policies, and those policies should address disclosure of data. A privacy policy represents an obligation to data subjects (whether they know they are data subjects or not), to data sources, and to the world. A privacy policy is a commitment that data will be used and disclosed in some ways and not others. In the EU and in other jurisdictions with broadly applicable privacy laws, researchers are subject to those laws and bound by their limitations. Whether we have an applicable law here or not, a promise made through a privacy policy should be strictly honored (and not subject to casual change by amendment of the privacy policy after-the-fact).[14]

---

[13] See Chris Jay Hoofnagle, *What Californians Understand about Privacy Online* (2008), at http://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID1262130_code364326.pdf?abstractid=1262130.

[14] No question was asked about changes to mandatory privacy policies. It is likely to be necessary to allow changes because the world is too dynamic to allow a fixed policy to remain unchanged under all circumstances. However, it is possible to allow changes to some parts of a policy without allowing everything to change. In crafting

Having said that, a privacy policy can still be ambiguous or discretionary with respect to disclosures for unforeseen research or other unanticipated events. It is unattractive to allow researchers wholly on their own to decide to pass PII to other researchers without some independent review. A neutral party should make judgments about the value of the secondary research, about the nexus of that research to the original project, about the importance of the data to the research objective, about alternative methods of obtaining useful data, and about whether the recipient's privacy and security policies are both adequate and consistent with the policies of the original data collector. It is hard to identify an existing institution other than an IRB to conduct this type of review. Without some assurance that a subsequent researcher will meet the same standards as the original researcher, there is a chance that there could be a daisy chain of data transfer from one researcher to another, with the shedding of data subject protections along the way. That possibility – including the prospect of data transfers to jurisdictions without privacy laws or with weaker privacy laws – must be addressed.

Here is a suggested policy that should be used to judge research data transfers without the consent of a research subject: A researcher can share data with another researcher if:

1) disclosure is expressly allowed by the disclosing research project's privacy policy;
2) the receiving research project meets the standards of the Common Rule;
3) the receiving research project has privacy and security policies that are as protective of privacy and security as the policies of the disclosing research project;
4) disclosure of the data has been approved by an IRB as consistent with these requirements, as scientifically valuable, and as compatible with the purpose of the transferring research project; and
5) the data transfers is memorialized in a data use agreement between the disclosing and receiving research projects.

HHS should prepare and mandate use of **model data use agreements** in order to make sure that they are complete and, in particular, that they strictly control use and redisclosure of data by the recipient. An independent third party, such as an IRB, should review all proposed transfers to assure compliance with these standards.

E. Consent and research on biospecimens collected outside of a research study

Question 47: Should there be a change to the current practice of allowing research on biospecimens that have been collected outside of a research study (i.e. ''left-over'' tissue following surgery) without consent, as long as the subject's identity is never disclosed to the investigator?

Response: This is a complex subject with no simple answer. However, the last condition suggested in the question is unrealistic. **Biospecimens with DNA cannot no longer be**

---

requirements for research privacy policies, HHS should pay attention to the terms under which changes are allowed and not allowed. For example, promises not to disclose data should not be subject to change without specific statutory direction.

considered non-identifiable. HHS is already well aware of the problem, and NIH recently took steps to withhold from public disclosure *aggregate* genetic information that no one previously thought could be individually identifiable.[15] While identifying a specimen today may still be difficult and expensive in many instances, the difficulty and expense of DNA analysis have been reduced enormously in the past decade. The trend is clear, and identifying an individual from DNA will be more possible, simpler to accomplish, and inexpensive in the near future. Nor is it hard to envision an accessible database of DNA to support identification in the near future, whether the database comes onto being as a result of expanded police activity, the implementation of a national health information infrastructure, more aggressive data collection by schools, private genetic testing outside the health care system, expanded newborn screening, or otherwise.

For HHS to make any policy based on the notion that a biospecimen is not identifiable would be a mistake that the Department would have to revisit soon. HHS should accept the inevitable, and proceed on the assumption that biospecimen identification will be both possible and practical. If that is the case, then conditions that accompany transfers of biospecimens or data from biospecimens should control activities by the recipient that might lead to identification of individuals. Data use (or specimen use) agreements are likely to offer the best way to establish the terms of sharing. This would be a step in providing suitable protections when specimens are used for research in ways not disclosed to or agreed to by the data subject. HHS could provide model agreements to help everyone adopt fair and comprehensive terms.

F. Consent of research subjects and meaningful opportunity to not consent

Question 50: What is the best method for providing individuals with a meaningful opportunity to choose not to consent to certain types of future research that might pose particular concerns for substantial numbers of research subjects beyond those presented by the usual research involving biospecimens? How should the consent categories that might be contained in the standardized consent form be defined (e.g. an option to say yes-or-no to future research in general, as well as a more specific option to say yes-or-no to certain specified types of research)? Should individuals have the option of identifying their own categories of research that they would either permit or disallow?

Response: It is only a matter of time before a research project comes to broad public attention because it involves an area of research that is ethically, religiously, morally, or politically objectionable to a portion of the American public. We predict that the trigger for public attention will be the use of data or biospecimens without individual consent. When that happens, legislative controls will be likely to follow (whether at the national or state levels) that will

[15] See Homer N, et al. (2008) *Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays*, PLoS Genet 4(8): e1000167 (2008), at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2516199. See also *The National Institutes of Health removes patients' genetic profiles from its website after a study reveals that a new type of analysis could confirm identities*, LA Times (Aug. 29, 2009), http://www.latimes.com/news/nationworld/nation/la-me-dna29-2008aug29,0,4364552.story.

mandate individual choice.[16] This is most likely to occur with health research, and these comments focus on health data. The same solutions will be needed elsewhere, but health is probably the most important area and the "easiest" to address.

HHS may as well accede to the inevitable now and give individuals the opportunity to make choices about the use of their own health data and specimens. Many people will agree to unrestricted use. Some will agree to use for some broad categories of research but not others. Some will want case-by-case consent. If supported by the health care record keeping system, others will subscribe to a research use template sponsored by organizations with particular points of view. HHS should allow for all of these options.

How to carry out this difficult task? For health data, at least, one way is to build the capability into the developing National Health Information Infrastructure, or Network (NHIN). HHS has spoken eloquently about using the NHIN to give individuals greater opportunities to participate in their own health care. Controlling the research use of their data is one way that people can participate. If choice were supported, the system could start with a default choice that allows for reasonable data use (under appropriate conditions) in recognition of the potential societal benefits of research. Those individuals with different requirements could change that default choice as they please.

Researchers will complain that any restrictions on full access will interfere with their scientific conclusions. Whether that is the case or not is more of an open question than researchers want to admit, but researchers are not the only ones with an interest here. Individuals who object to the use of their data and their biospecimens have an interest that must be considered and accommodated in a reasonable way.

G. HIPAA identifiability standards

Question 54: Will use of the HIPAA Privacy Rule's standards for identifiable and de-identified information, and limited data sets, facilitate the implementation of the data security and information protection provisions being considered? Are the HIPAA standards, which were designed for dealing with health information, appropriate for use in all types of research studies, including social and behavioral research? If the HIPAA standards are not appropriate for all studies, what standards would be more appropriate?

Response: As HHS knows, the HIPAA identifiability standards have been questioned and are under review internally. How those standards may change is unknown. Standards for identifiability must change over time as technology changes and as more personal data is collected, compiled, and made available. The realm of truly de-identifiable data will continue to shrink over time, and there may be little de-identified data anywhere in the foreseeable future.

[16] The well-publicized example involving the Havasupai Indians offers a taste of the controversy that can arise. See, e.g., Amy Harmon, *Indian Tribe Wins Fight to Limit Research of Its DNA*, New York Times (April 21, 2010), at http://www.nytimes.com/2010/04/22/us/22dna.html.

The narrow answer is that the current HIPAA identifiability standards may not be adequate for their purpose or for research. However, it would be appropriate for the Common Rule to use the same identifiability standard used for HIPAA as much as possible. Indeed, it would be a poor policy if there were multiple standards of identifiability, and researchers had to figure out which one applied to their projects. Policy shopping might be an unfortunate possibility.

Unfortunately, more than one standard of identifiability may be inevitable. The HIPAA standard is finely tuned to health records. If applied to the wide variety of records used in other research activities, the HIPAA standard might not work. A different standard might apply to biospecimens. Other types of research will require their own standards. Identifiability is a continuum, and there are rarely assurances that de-identified data is truly anonymized. Finding the "sweet spot" for any type of data and research requires judgment.

One way to minimize the multiple standards problem is to establish a policy that limits conflict over which standard applies under which circumstances. Putting anyone is a situation where two (or more) overlapping identifiability standards both apply would be most unfortunate. Thus, the HIPAA standard could apply in full and exclusively to any research project that has health data covered by that standard. A rule that applies one and only one identifiability standard to a research project is sensible as long as the applicable standard is the strictest. Other problems of identifiability will have to be addressed in other ways, including broad use of data use agreements.

One proposal for greater use of data use agreements can be found in a recent article.[17] The abstract summarizes the idea:

> Deidentification is one method for protecting privacy while permitting other uses of personal information. However, deidentified data is often still capable of being reidentified. The main purpose of this article is to offer a legislative-based contractual solution for the sharing of deidentified personal information while providing protections for privacy. The legislative framework allows a data discloser and a data recipient to enter into a voluntary contract that defines responsibilities and offers remedies to aggrieved individuals.

While the article calls for legislation, some of the proposed goals could be accomplished through regulations that rely upon data use agreements as a mandatory method for balancing the interests of individuals and the needs of researchers by allowing some transfers of data while establishing standards for the transfers and liability for those providing and receiving data.

H. DNA from de-identified biospecimens and identifiability

Question 56: DNA extracted from de-identified biospecimens can be sequenced and analyzed in other ways, with the results sometimes being linked to other available data than may allow a

[17] Robert Gellman, *The Deidentification Dilemma: A Legislative and Contractual Proposal*, 21 Fordham Intellectual Property, Media & Entertainment Law Journal 33 (2010), at http://iplj.net/blog/wp-content/uploads/2010/11/C02_Gellman_010411_Final.pdf.

researcher to identify the persons whose specimens were being studied. How should Federal regulations manage the risks associated with the possibility of identification of such biospecimens? Should a human biospecimen be considered identifiable in and of itself? What are the advantages and disadvantages of considering all future research with biospecimens to be research with identifiable information?

Response: A point made in responding to question 47 applies here. Biospecimens with DNA cannot be considered non-identifiable. HHS must accept this make policy accordingly. Data use agreements should be required to establish the regulatory standards that apply to the transfer of identifiable or potentially identifiable data or specimens. The proposal for universal privacy and security standards will also help in this context, but privacy and security standards will not address all the issues involved.

I. Applicability of rules?

Question 58: Should the new data security and information protection standards apply not just prospectively to data and biospecimens that are collected after the implementation of new rules, but instead to all data and biospecimens? Would the administrative burden of applying the rule to all data and biospecimens be substantially greater than applying it only prospectively to newly collected information and biospecimens? How should the new standards be enforced?

Response: It would be administratively impossible for most organizations to have privacy and security standards that apply to some research data but not to others. HIPAA adopted a policy that essentially applied the same requirements to all health care data, past, present, and future. Whatever problems a retrospective policy for information will encounter will be easier to deal with than a prospective-only policy. A prospective policy will create so many challenges and conflicts that it will tie lawyers into knots, resulting in different applications to different data by different institutions. The resulting mess will be unintelligible to nearly everyone. Application of different policies to different data collected at different times will be doomed to failure. There may be a need for transition rules for some data or activities, but there should be a single policy applicable to all past, present, and future data.

The treatment of specimens may justify some differences between specimens collected in the past and specimens collected after new policies are adopted. One reason is that it may be easier to identify and segregate specimens by the time of collection. The same is less likely to true for information, especially when the same information is collected under two (or more) different privacy regimes.

J. Standards and protections

Question 59: Would study subjects be sufficiently protected from informational risks if investigators are required to adhere to a strict set of data security and information protection standards modeled on the HIPAA Rules? Are such standards appropriate not just for studies involving health information, but for all types of studies, including social and behavioral research? Or might a better system employ different standards for different types of research? (We note that the HIPAA Rules would allow subjects to authorize researchers to disclose the

subjects' identities, in circumstances where investigators wish to publicly recognize their subjects in published reports, and the subjects appreciate that recognition.)

Response: The notion of a single set of privacy and security standards based on HIPAA is attractive, but it is unworkable and bad policy in some ways with respect to the HIPAA privacy rule but not with respect to the HIPAA security rule.

The HIPAA privacy rule is finely attuned to the health care system. While some of the provisions of the privacy rule would work well in other contexts (e.g., notice to data subjects of subpoenas), some would not. The clearest (but not the only) examples of HIPAA privacy policies that will not work elsewhere come from the privacy rule's expansive disclosure provisions.

The most controversial of these provisions allows 1) disclosure of all health care records to law enforcement with standards to weak that they do not even require a written request for the records; and 2) disclosure of all health care records to an intelligence agency without any standard at all and without even a request by an intelligence agency. The inappropriateness of these disclosure policies for health records is beyond the scope of these comments. However, there is no reason why these rules (and those for disclosure to coroners, prisons, families, etc.) should apply in general to health research or more broadly to all research subject to the Common Rule.

The political pressures that forced HHS to adopt a HIPAA privacy rule that allowed largely untrammeled access to law enforcement and intelligence agencies may not recur in the research context. It is possible that a narrower disclosure rule could make it through the regulatory process without a reaction from law enforcement and intelligence agencies. However, if that judgment is wrong and there is a political reaction, a mandated policy that allowed all researchers to make their records available to police with minimal procedure and to the CIA with no procedure would create an issue that would plague researchers forever. Health care is not optional, and few people pay attention to the breadth of the HIPAA disclosures. That would be less so in the case of research, and anyone potential research subject who actually read a disclosure policy allowing CIA disclosure would be much more likely to look for the nearest exit.

The best answer for privacy is to mandate application of the classic standards of Fair Information Practice. Implementation of those standards may vary with the nature of the data and the type of institution involved. Most people would agree that a stricter application of FIPs is more appropriate for hospital records than for pizza delivery records. Both types of records will be the subject of some types of research, with the right result being differing implementations of the same broad FIPs policies.

A way to avoid the political problem of law enforcement and intelligence agency access is to leave that part of the policy to local determination. A mandated policy should address disclosure, of course. A revised Common Rule might address standard categories of disclosure, such as use for internal activities directly related to the research (employees and volunteers) and for external activities in direct support of the research (lawyers, auditors, and service providers). A mandated policy should prohibit marketing disclosures. For other disclosures, including law enforcement

and intelligence, researchers themselves might determine whether, when, and how to allow the disclosures subject to IRB review. In this way, rather than HHS imposing weak privacy policies, the decision can be left to each researcher. Other, ethically complex disclosures, such as those to individuals at risk, can be left to local determination and IRB review as well.

While the HIPAA privacy rule is not appropriate for extension to all research or even health research, the HIPAA security rule is different. The security rule is much less tied to the institutions holding data or the nature of the data being protected. The rule, with its required and addressable parts, is more likely to work generally. It would certainly be undesirable to ask institutions to comply with multiple security standards if that result can be avoided. The addressable parts of the HIPAA security rule, which allow record keepers to make local determinations about whether or how to apply some security elements, strike a reasonable balance. If a security rule other than the HIPAA rule is found necessary, then any institution covered by the HIPAA security rule (or comparable alternate security rule) should be left to comply with a single security rule only. Allowing the application of multiple security rules would be confusing, expensive, and unnecessary.

K. Additional standards beyond HIPAA?

Question 61: Are there additional data security and information protection standards that should be considered? Should such mandatory standards be modeled on those used by the Federal government (for instance, the National Institute of Standards and Technology recently issued a ''Guide to Protecting the Confidentiality of Personally Identifiable Information.'')?

Response: The basic idea of relying on existing models is a good one. However, the NIST document suggested is not a good example. By its own terms, the NIST recommendations "are intended primarily for U.S. Federal government agencies and those who conduct business on behalf of the agencies." Federal agencies and some federal contractors are subject to the Privacy Act of 1974, an old, largely outdated privacy law that cannot be readily applied in other contexts. The Act's use of a *system of records* framework and its expansive disclosure authority make it inappropriate for broader application. Despite the Act's limited utility, Fair Information Practices (FIPS) are at the heart of the Act. FIPs should be used as the organizing policy for privacy under a revised Common Rule.[18]

L. Disclosure and data use agreements

Question 62: If investigators are subject to data security and information protection requirements modeled on the HIPAA Rules, is it then acceptable for HIPAA covered entities to disclose limited data sets to investigators for research purposes without obtaining data use agreements?

Response: No. Indeed, if investigators are subject to HIPAA-like rules, then the need for data use agreements will be even greater. HIPAA allows too many disclosures that are inappropriate for researchers. Further, different research projects may have different policies, and there will be

[18] For a history of FIPs, see Robert Gellman, Fair Information Practices: A Basic History, at http://bobgellman.com/rg-files/rg-resources.html.

a need to reconcile policies when sharing data. Data use agreements would be the best way for HIPAA covered entities to limit their own legal exposure by restricting how researchers can use and disclose data and meet other privacy requirements. HHS can help by commissioning a series of **model data use agreements** for different contexts. This will support data transfers under reasonable terms without requiring everyone to expend resources needlessly on hiring lawyers and reinventing the wheel.

If HHS contemplates a single, mandated-from-Washington privacy and security policy for all researchers, then this answer might need modification, depending on just what that policy is. However, a much better approach would be to give *some* discretion to researchers in adopting their own privacy and security policies because local requirements will vary considerably. Whether local policies are sufficient may require more of a case-by-case review by those who are supplying data to researchers. How could anyone expect a data source to accept the liability that goes along with transferring data to a researcher without doing some due diligence? The best part is that a review of researcher policies by data sources will oblige researchers to adopt sufficiently protective policies to satisfy those sources. This will help to make the system self-correcting, as researchers will strive to establish their own policies that will satisfy data sources without complication. Common standards and practices will develop over time, with benefits to data subjects, researchers, and data sources.

M. Re-identification issues

Question 63: Given the concerns raised by some that even with the removal of the 18 HIPAA identifiers, re-identification of de-identified datasets is possible, should there be an absolute prohibition against re-identifying de- identified data?

Response: It is not clear how the Common Rule would be able to accomplish an absolute prohibition even if it wanted to. At some level, de-identified data is likely to pass beyond the scope of regulation, as is the case under HIPAA. If so, anyone may acquire the data and seek to re-identify it. Many of those potential re-identifiers will not be subject to the Common Rule or the jurisdiction of HHS enforcement.

Further, there may well be some types of research for which re-identification is essential, even if only to determine the extent to which re-identification is possible. A flat ban on re-identification, even if accomplished by a new statute, is probably too crude an instrument. In at least some (narrow and hopefully well-defined) contexts, re-identification may be reasonable. What is most in need of control may not be re-identification per se, but how any re-identified data can be used. The individual whose data is re-identified may not need protection against the statistical researcher who discovers a new re-identification technique and publishes an article about the technique in a scholarly journal. The individual need protection against the banker, employer, insurer, and others who may seek to use re-identified data in ways that will affect the individual's interests.

Re-identification needs to be controlled through data use agreements, and preferably through agreements that include overt **third-party beneficiary clauses** that would allow aggrieved data subjects to sue anyone who re-identified their data in violation of a data use agreement.[19]

N. Disclosure of de-identified data and prohibitions

Question 64: For research involving de-identified data, is the proposed prohibition against a researcher re- identifying such data a sufficient protection, or should there be in some instances requirements preventing the researcher from disclosing the de-identified data to, for example, third parties who might not be subject to these rules?

Response: See the answer to question 63. There must be data use agreements covering all transfers because there is no other mechanism with the potential to bind recipients in the absence of legislation. Frankly, even legislation will not suffice because data can be easily transferred to other jurisdictions beyond the reach of U.S. law.

O. Reporting of data?

Question 68: With regard to data reported to the Federal government:

a. Should the number of research participants in Federally funded human subjects research be reported (either to funding agencies or to a central authority)? If so, how?

b. What additional data, not currently being collected, about participants in human subjects research should be systematically collected in order to provide an empirically-based assessment of the risks of particular areas of research or of human subjects research more globally?

Response: It is not clear precisely what is contemplated in these questions. However, one answer is clear. A central database with identifiable information about participants in human subjects research is a terrible idea. This is true regardless of the purpose of the database. If a database of that type is contemplated here, it should be dropped as a possibility.

Any new database with information about individuals will become a magnet for users, including many who are not specifically foreseeable today. The principle is *if you build it, they will come*. For example, a database will eventually attract those involved in law enforcement, immigration control, child support enforcement, and more. It would be too easy for the Congress to enact broad legislation in these and other areas mandating that all information resources be enlisted in support of whatever the issue of the day happens to be. That has happened already. In addition, a central database will create enormous security risks, and the inevitable security breach will impose huge costs on its sponsor and significant problems for data subjects.

[19] A brief discussion of the issue of third-party beneficiaries can be found at Robert Gellman, *The Deidentification Dilemma: A Legislative and Contractual Proposal*, 21 Fordham Intellectual Property, Media & Entertainment Law Journal 33, 50 (2010), at http://iplj.net/blog/wp-content/uploads/2010/11/C02_Gellman_010411_Final.pdf

The collection of aggregate statistics that are not identifiable and are not re-identifiable raises no privacy concerns. We note that the pool of genuinely non-identifiable data is shrinking, and will continue to do so. Whether statistical or other, central reporting as a worthy idea is not addressed in this response.

Respectfully submitted,

s/

Pam Dixon
Executive Director,
World Privacy Forum