



October 21, 2011

Jerry Menikoff, M.D., J.D.
Office for Human Research Protections
1101 Wootton Parkway, Suite 200
Rockville, Maryland 20852

***Re: Comments in response to the Advanced Notice of Proposed Rule Making (ANPRM) on
“Human Subjects Research Protections”***

Dear Dr. Menikoff:

The National Committee on Vital and Health Statistics (NCVHS) is the statutory public advisory body on health data, statistics, and national health information policy to the Department of Health and Human Services (HHS). NCVHS is submitting these comments in response to the Advanced Notice of Proposed Rule Making (ANPRM) on “Human Subjects Research Protections: Enhancing Protections for Research Subjects and Reducing Burden, Delay, and Ambiguity for Investigators.”¹ These comments are directed to topics in the ANPRM on which NCVHS has previously conducted hearings or prepared reports and are based on these published materials. We begin with some general points and follow with answers to selected specific questions posed in the ANPRM on which NCVHS has particular expertise.

Introduction

The ANPRM proposes to reform regulations governing research with human subjects in an attempt to both improve the review process and enhance protection of research subjects. The proposed approach includes far-reaching changes in the Common Rule, some of which may require conforming changes to the Health Insurance Portability and Accountability Act regulations (HIPAA Privacy and Security Rules). One type of change involves re-evaluation of risks to individuals relating to their information and a proposal for a new structure to govern these risks. Another type of change involves a broad set of efforts to streamline reviews of research. As the statutory advisory body on health data, NCVHS has extensive experience with the protection of information and with the importance of the use of data in research involving public health, health care, and the improvement of health.

A basic premise of the ANPRM is that “most research risks to the individual can be categorized into one of three types: physical, psychological, and informational risks.” The ANPRM states that harms such as “legal, social, and economic” harms can “usually be viewed as variations on those core categories.”² Further, “[i]nformational risks derive from inappropriate use or disclosure of information, which could be harmful to the study subjects or groups.” In the judgment of the ANPRM, “the majority” of unauthorized disclosures of identifiable health

¹86 Fed. Reg. 44512 (July 26, 2011).

²45 Fed. Reg. 44515 (July 26, 2011).

information used in research are the result of inadequate data security, that is, of inappropriate release of the information.³ Relying on these assumptions, the ANPRM recommends a new approach to protecting research subjects against information risks, placing such risks largely outside the purview of Institutional Review Board review and instead within the protections of the HIPAA Privacy and Security Rules. However, some research may be carried out by entities not covered by HIPAA and thus evade review or regulation entirely if regulation is to be based solely on HIPAA.

NCVHS' prior work on data stewardship, secondary uses of data, data privacy and data security gives good reasons to question these assumptions and approach. In addition to disclosures of information, uses of information may also be troubling from a privacy perspective. Moreover, much of the data about human subjects used in research does not come within the HIPAA Privacy and Security Rules as currently structured and may not fit well within these rules. NCVHS' long history of support for the collection of data needed for research concerning population health and the health and health care of underserved groups also suggests reasons for concern about the approach of the ANPRM with respect to the importance of community trust in data use.⁴ Based on its prior work, NCVHS believes that legal, social, economic, and cultural harms cannot be reduced in simple fashion to informational harms, and that security risks relating to inappropriate release of information, although certainly important, should not be singled out as the primary risks warranting protection. Even when security is protected, these other risks are not addressed.

In our publication "Health Data Stewardship: What, Why, Who, How: an NCVHS Primer,"⁵ NCVHS emphasized the importance of data stewardship practices for all those, including researchers, who use individual health data. These stewardship practices, NCVHS explained, must "strengthen[] the chain of trust and accountability" so that data may be used to the greatest possible benefit while risks of harm are minimized. This Primer built on our April 2008 Report, "Enhancing Protections for Uses of Health Data: A Stewardship Framework,"⁶ in which NCVHS described how privacy and confidentiality protection and data accountability, as well as security, are critical aspects of stewardship. In our initial stewardship report, "Enhanced Protections for Uses of Health Data: A Stewardship Framework for 'Secondary Uses' of Electronically Collected and Transmitted Health Data," NCVHS emphasized the importance of data stewardship practices when data originally collected in patient care are used for other purposes ("secondary uses"). Research is an important such secondary use, as is public health. Harms that may result from data uses that individuals do not expect and have not consented to include loss of trust in the health care system and resulting refusals of patients to share information about their health, potentially risking their own health and the public health. Harms

³45 Fed. Reg. 44516 (July 26, 2011). This assumption that informational harms are primarily traceable to inappropriate release of information is also present in a summary of the ANPRM published in the *New England Journal of Medicine*, Ezekiel J. Emanuel & Jerry Menikoff, Reforming the Regulations Governing Research with Human Subjects, *NEJM* 364(12): 1145-1150 ("Many studies pose risks that are largely informational: harms result primarily from the inappropriate release of information and not from the research interventions themselves.") However, it is not clear on what evidence this assumption is based.

⁴ (See, e.g., [Letter to the National Institute of Child Health and Human Development, NIH](#) regarding the establishment of the National Children's Study, February 20, 2004, <http://www.ncvhs.hhs.gov/040220lt.htm>; [Report to the Secretary - Eliminating Health Disparities: Strengthening Data on Race, Ethnicity, and Primary Language in the U.S](#), November 7, 2005, <http://www.ncvhs.hhs.gov/051107rpt.pdf>).

⁵<http://www.ncvhs.hhs.gov/090930lt.pdf> (September 2009).

⁶<http://www.ncvhs.hhs.gov/090930lt.pdf> (April 24, 2008).

National Committee on Vital and Health Statistics

also may include discrimination, stigmatization, personal embarrassment, and harms to groups and their values.

At the same time, NCVHS recognizes the importance of data use in research to improve public and community health, health care quality, and effectiveness of medical treatment. NCVHS therefore applauds many of the ways in which the ANPRM would streamline and harmonize reviews of the uses of health information in research and for public health.

NCVHS believes the goal of the Common Rule should be to protect individuals and data about individuals in each setting. We fully agree with the goals of the ANPRM to better protect human subjects while facilitating research and reducing ambiguity and investigator burden. The protection of the public's health relies on our ability to collect and use information for myriad public health purposes. When we do so, however, we must respect individual privacy and guarantee ethical conduct by researchers. The current system of IRB review needs improvement, but in our effort to become more efficient, we must not diminish the essential role and protections this system provides.

NCVHS thanks HHS for the opportunity to comment on the ANPRM and looks forward to working with the Department as it moves forward to improve the protection of research subjects and the ability of researchers to do their important work.

Respectfully submitted,

/s/

Justine M.Carr, M.D.
Chairperson,
National Committee on Vital and Health Statistics



Answers to Specific Questions Posed in the ANPRM

In what follows, NCVHS utilizes our general observations about the different types of information risks, the importance of trust, and the benefits of data use to answer selected questions posed in the ANPRM. Questions are listed by their assigned number in the ANPRM.

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?

This question poses two different issues. The first issue is whether research (including survey research) collecting certain types of information should be regarded as more than minimal risk.

NCVHS has documented that certain types of information regarded as especially sensitive have been given special protections under a variety of state and federal laws. Under federal law, these types include genetic information, psychotherapy notes, substance abuse treatment records, and treatment paid for fully in cash where the patient requests the payer not know. Under state laws, the types of information include HIV status or sexually transmitted diseases, mental health information, information in the records of children and adolescents, and sexuality and reproductive health information. By enacting these laws for specially protected information that include heightened de-identification standards, and restrict permissible uses and disclosures, states made policy decisions that use and disclosure of certain categories of information pose heightened risk to their citizens. Congress recognized states' rights to make these decisions when it enacted HIPAA and provided that more stringent state laws shall not be preempted. Moreover, in certain contexts additional categories of information may be sensitive, such as mental health, sexuality, reproductive health, adolescence, and domestic violence information; NCVHS has identified the protection of these types of information also as being important to preserving patient trust.⁷

Under the Common Rule, research is “minimal risk” if “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.”⁸ NCVHS findings regarding these categories of sensitive information indicate that there are contexts within which information in these categories might pose risks to individuals beyond those meeting this definition. For example, the probability or magnitude of the economic harm of loss of a job or of insurance coverage that Congress attempted to protect against by laws such as the federal Genetic Information Non-Discrimination Act might be beyond that which would be ordinarily encountered in daily life. However, many such laws have been criticized as not wholly effective. Thus NCVHS believes studies involving these types of sensitive information that have been singled out for special legal protection could be more than minimal risk in some contexts. Likewise, information not specifically protected by law but the knowledge of which could invite political controversy, stigmatization, discrimination that is not legally prohibited (such as against homosexuals), speculation, harassment, or undesired attention could also be more than minimal risk in some contexts.

⁷Letter to the Secretary, Recommendations Regarding Sensitive Health Information (Nov. 10, 2010), <http://www.ncvhs.hhs.gov/101110lt.pdf>.

⁸45 C.F.R. § 46.102(i) (2011).

The second issue raised by this question is whether the characteristics of particular populations should be taken into account in making the risk assessment. There are several reasons why this might be so: the population at issue faces risks that are different from the risks faced by the population at large, the population at issue is more vulnerable to certain risks than the population at large, or the population at issue is less able to protect itself from risks through mechanisms such as informed consent. With respect to difference in risks and vulnerability to them, NCVHS has long documented the differential health risks faced by different populations. In these comments, we do not address the ability of particular populations to protect themselves effectively through informed consent or other mechanisms, although we recognize that this is an issue as well.

14. Are these expansions in the types of studies that would qualify for this Excused category appropriate? Would these changes be likely to discourage individuals from participating in research? Might these changes result in inappropriately reduced protections for research subjects, or diminished attention to the principles of respect for persons, beneficence, and justice?

This question refers to the proposal to create a new category of “excused” research that would not require IRB review, but would require certain more limited safeguards. The question is highly general and answering it requires further understanding of the details of what is proposed. This category would replace and expand the current categories of “exempt” research. “Excused” research would include all educational tests, surveys, focus groups, interviews, and similar procedures when the subjects are competent adults. Additionally, it is possible that a mechanism might be created to add certain “benign interventions” using common social science techniques, such as word association or time performance tests, to the “excused” category. Secondary uses in research of data or biospecimens originally collected in clinical care, even with identifiers, would also be “excused,” unless there are plans to provide research results back to subjects.⁹

For such pre-existing data or biospecimens collected in clinical care, new, streamlined consent requirements also are envisioned. A standardized general consent form is proposed that could be given at the time of original collection of the data or biospecimens. This form is expected to permit the subject to accept or refuse participation in all future research, or to specify a “handful” of special categories of research with biospecimens that could be refused. For pre-existing data collected for clinical care, written consent would only be required for data with identifiers. For pre-existing data collected in the first instance for research, written consent would be required whether or not the data contain identifiers. All biospecimens would be understood to contain identifiers, and subsequent use of these specimens in research would require written consent.

In responding to this question, NCVHS believes that it is important to clarify what is meant by “data with identifiers.” In HIPAA, data are considered de-identified if 17 specified data types as well as “[a]ny other unique identifying number, characteristic or code” have been stripped from the data.¹⁰ However, other laws define identifiers differently, and there are a number of different methodologies for avoiding re-identification.

⁹ 76 Fed. Reg. 44519 (July 26, 2011).

¹⁰ 45 C.F.R. § 165.514 (b)(2) (2011).

National Committee on Vital and Health Statistics

In early recommendations concerning HIPAA, NCVHS emphasized the importance of either authorization or waiver of authorization by an entity knowledgeable about privacy protection for secondary uses of information in research.¹¹ NCVHS remains concerned about the privacy implications of open-ended consent, especially where sensitive types of information or research are in question. In this connection, it is important to bear in mind the types of information risks described above that extend beyond security risks.

These concerns extend to public health research. Public health research relies heavily on the use of secondary data. These data come from clinical and laboratory records, public health surveillance systems such as vital statistics and disease registries, and a variety of other sources, some of which are covered by HIPAA and others of which are not. Thus only some of this data currently receives HIPAA protection and the data sets employed may be made up of both data originally protected by HIPAA and data that are not protected by HIPAA. Moreover, many patients, clients, and other individuals represented in these data systems are unaware that this information may be used for research. Although informational risks might be low, they do exist as discussed above. Ethical concerns require that IRBs be actively involved in the determination of whether and what type of informed consent is required prior to the use of these data. NCVHS is concerned that exempting or excusing the use of secondary data for public health research from IRB review poses potential informational risks to the human subjects.

Finally, NCVHS registers the concern that the use of the term “Excused” suggests that studies would be out of the purview of any protections. Although we would prefer a different term, the term “registered” (the alternative suggested in Question 20), makes sense only if HHS adopts the registration and audit proposal detailed in the ANPRM. If HHS does not adopt those proposals, we would prefer that any future NPRM use a term that better captures the need for oversight of this research.

16. Should research involving surveys and related methodologies qualify for the Excused category only if they do not involve topics that are emotionally charged, such as sexual or physical abuse? If so, what entity should be responsible for determining whether a topic is or is not emotionally charged?

As an alternative to the subjective term “emotionally charged,” NCVHS has been guided by state and federal law to identify categories of sensitive information that may require separate management. Information falling into these categories may raise special concerns. These laws are based on factors beyond whether the information is “emotionally charged,” including possible risks to patients of the use of this information. Risks of genetic information, for example, may involve not only whether the information is “emotionally charged” but whether the information may be the subject of employment or insurance discrimination, or whether the information may be critical to group identity and any potential associated stigma.

In our letters concerning sensitive information, NCVHS has also explained the difficulties for patient care and data management if individuals decide which data elements, at too granular a level, they regard as personally sensitive. Instead, we suggested the determination

¹¹ 2001 letter on HIPAA and research: importance of authorization or waiver.(NCVHS Letter to the Secretary of HHS, Nov. 21, 2001, <http://www.ncvhs.hhs.gov/011121lt.htm>).

of categories of information that should be identified for special management, as laws require.¹² This approach would suggest the incorporation of categories of information or research that are regarded as sensitive into regulations such as a newly-structured Common Rule, rather than determination by individual patients or IRBs case by case. Such categorization could be applied to the consent process involving the future use of information or biospecimens collected in clinical care, for example. In the general consent forms envisioned by the ANPRM, patients could be offered the opportunity to consent to all research involving their information or biospecimens, or to consent to no research, or to consent to all research except that falling into the specified categories.

NCVHS' hearings and report on sensitive categories of health information, referred to above, suggest that some patients might be concerned about the use of particular types of their health information in research. We have received testimony that patients who could not be assured separate management of information in these sensitive categories might withhold information in the context of clinical care, with potentially adverse consequences to their health or the public's health. This suggests the possibility that it is not only types of research but also types of information that might require separate management for research purposes. In this connection, NCVHS notes that ONC is currently piloting a data segmentation initiative for separate management of specified categories of sensitive information.¹³

24. The Common Rule has been criticized for inappropriately being applied to—and inhibiting research in—certain activities, including quality improvement, public health activities, and program evaluation studies . . . We seek comment on whether and, if so, how, the Common Rule should be changed to clarify whether or not oversight of quality improvement, program evaluation studies, or public health activities are covered . . .

With respect to quality improvement, program evaluation, and public health, this question asks whether the existing rules—even with the proposed changes—are inappropriate, whether the problem should be addressed through definition of “excused” categories or through re-definition of “research,” or whether the protections of the Common Rule continue to be necessary. The question addresses three categories together that are handled differently in the HIPAA Privacy Rule. This is one area where the Common Rule and the HIPAA Privacy Rule are not aligned. Under the HIPAA Privacy Rule, quality improvement and program evaluation fall into the category of “health care operations,” but public health activities do not. Therefore, some public health activities may be regarded as disclosures for the purpose of research under the HIPAA Rules while quality improvement and program evaluation are not. In this comment, NCVHS addresses the use of information in research related to public health activities.

The use of health information for public health research and the improvement of public health is an important public good and a critical aspect of the learning health care system envisioned by the HITECH Act. The HIPAA Privacy Rule allows the transfer of protected health information without authorization to a public health authority authorized by law to collect

¹²Letter to the Secretary, Individual control of sensitive health information accessible via the Nationwide Health Information Network for purposes of treatment (Feb. 20, 2008), <http://www.ncvhs.hhs.gov/080220lt.pdf>; Letter to the Secretary, Recommendations Regarding Sensitive Health Information (Nov. 10, 2010), <http://www.ncvhs.hhs.gov/101110lt.pdf>.

¹³See <http://www.healthit.gov/buzz-blog/from-the-onc-desk/announcing-metadata-pilots-realize-pcast-vision/>.

National Committee on Vital and Health Statistics

or receive it for a public health purpose.¹⁴ Public health authorities are mandated to use the information to improve public health, including through research. With respect to data collection, public health authorities are not included in the HIPAA definition of business associates, and, therefore, they are not required to have business associate agreements with covered entities from whom they receive health information, or with their public health partners or contractors.¹⁵ In fact, public health authorities collect protected health information (PHI) from covered entities and others to fulfill their statutorily mandated missions and do not perform these collection functions on behalf of any covered entity. State or federal law and regulations provide not only the authority for the data collection, but also the structure for privacy, security, and governance of activities involving the data.

The line between public health activities such as surveillance and disease prevention, and research, may be difficult to draw, however, as activities initially begun as public health measures may shift into the collection and analysis of data intended to create generalizable knowledge.¹⁶ Members of the public are unlikely to distinguish between public health improvement and research in how their data are protected. NCVHS' recent hearings concerning community uses of data have underlined the importance of the use of data as well as the need for trust and a common set of protections. Rather than individualized informed consent models such as transparency and opt-out capabilities, community engagement in setting research priorities, and community involvement in sharing research results may be critical methods for engendering trust.

31. How does local IRB review of research add to the protection of human subjects in multi-site research studies? How would mandating one IRB of record impair consideration of valuable local knowledge that enhances protection of human subjects? Should the public be concerned that a centralized IRB may not have adequate knowledge of an institution's specific perspective or the needs of their population, or that a centralized IRB may not share an institution's views or interpretations on certain ethical issues?

Although the proposal to designate a single IRB of record has the appeal of greater efficiency, there are significant practical concerns for implementation. In particular, public health research data coming from differing state government data bases, such as state vital registries, may be subject to a variety of state laws, policies, and procedures that must be addressed. For example, one jurisdiction may require that the state health department make the first contact to persons in the state cancer registry to determine if they wish to be contacted for a multi-state research study. In this case, having multiple IRB reviews could be a necessary human research protection rather than an impediment to efficient research. Alternatively, the designated IRB could be required or encouraged to seek advice or other input from states in which a study operates to ensure compliance with its laws. Coordination of IRB review is a laudable goal; standards for selection of one designated or lead IRB of record need thoughtful consideration, however.

¹⁴45 C.F.R. § 164.512(b) (2011).

¹⁵ HIPAA Privacy Rule and Public Health: Guidance from CDC and the Department of Health and Human Services (April 11, 2003), <http://www.cdc.gov/mmwr/preview/mmwrhtml/m2e411a1.htm>.

¹⁶ This is the definition of research under the Common Rule, 45 C.F.R. § 46.102(d) (2011).

National Committee on Vital and Health Statistics

39. *If changes are made to the informed consent requirements of the Common Rule, would any conforming changes need to be made to the authorization requirements of the HIPAA Privacy Rule?*

Yes, there would need to be conforming changes to the HIPAA Privacy Rule. NCVHS notes, however, that there may also be a need to address gaps in protection created by the changes in the Common Rule.

For purposes of this question, the crucial differences between the Common Rule and the HIPAA Privacy Rule concern the concept of a limited data set and the scope of authorization. With respect to the requirement of authorization, HIPAA draws critical distinctions between de-identified data (by HIPAA standards), limited data sets, and identifiable PHI. Authorization is required for the last but not for either de-identified data (which are considered to be outside the purview of HIPAA) or limited data sets. The current interpretation of the HIPAA Privacy Rule as it applies to authorization for the use of identifiable PHI in research is quite stringent, requiring that the authorization be study-specific. Like HIPAA, the Common Rule applies to research involving human subjects; if research does not involve contact with individuals or “identifiable private information,” it is not considered to be research involving human subjects.¹⁷ Any identifiable information (which may include limited data sets) requires either informed consent or a waiver of informed consent granted by an IRB. Waivers must meet conditions specified in the Common Rule, including that it be impracticable to conduct the research without the waiver.¹⁸ The Common Rule will in some contexts allow consent to general types of research (e.g. cancer studies) rather than study-specific consent. This “misfit” between the HIPAA Privacy Rule and the Common Rule has been the subject of significant discussion and request for public input by HHS.¹⁹ This issue clearly needs to be settled, but in the judgment of NCVHS, the ANPRM as proposed does not address the problem adequately.

The approach proposed in the ANPRM, of allowing highly general consents to the use of identifiable information originally collected for clinical care (including biospecimens), may exacerbate the discontinuity between the HIPAA requirement for study-specific authorization and the Common Rule. So may the possibility that biospecimens could continue to be considered de-identified under HIPAA standards, thus not requiring authorization, even though they are considered identifiable under the ANPRM approach. The ANPRM’s proposed requirement for consent for the re-use of information originally collected for research, even if the information has been de-identified, also will impose a new requirement where HIPAA does not now apply. This is because information that has been de-identified in accordance with the HIPAA standard no longer qualifies as PHI and is thus outside the HIPAA regime. On the other hand, the proposed approach of the ANPRM will weaken the current existing protections for information collected for clinical care that does not come within HIPAA because it is not possessed by a covered entity—that is, the protection of the Common Rule would now apply in such cases.

¹⁷ The Common Rule definition of “human subject” is “a living individual about whom an investigator (whether professional or student) conducting research obtains (1) Data through intervention or interaction with the individual, or (2) Identifiable private information” 45 C.F.R. § 46.102(f).

¹⁸ 45 C.F.R. § 46.116(c)(2) (2011).

¹⁹ ANPRM, 76 Fed. Reg. 44523 (July 26, 2011).

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 . . . be defined? . . . Should individuals have the option of identifying their own categories of research that they would either permit or disallow?

In our sensitive information letters, NCVHS urged that interoperable health records be developed with the capacity for separate management of defined categories of sensitive health information.²⁰ In addition to options permitting all research and permitting no research, general consent to prospective uses of data could set out defined categories of information or research that would permit separate opt out possibilities for patients to choose. As explained above, NCVHS has documented confusion that might arise if patients defined their own types of sensitive information on a case by case basis. NCVHS therefore suggests specifying standard definitions of sensitive categories of information for which patients might choose separate management for research purposes.

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?

The first question asked here is whether the HIPAA standards for identifiable information, de-identified information, and limited data sets will facilitate implementation of data security and information protection provisions. Two of the major components of informational risk are loss of control of data and unauthorized re-identification. Unauthorized re-identification could occur using data disclosed for the purpose of research as the source, or using reported results as the source. It is important to distinguish that the HIPAA de-identification standard only applies to the former category and not to reported results. Once de-identified information is disclosed, it is no longer regulated by HIPAA at all. In the judgment of NCVHS, these risks remain inadequately addressed by the HIPAA standards.

NCVHS has heard testimony questioning the continued utility of the HIPAA safe harbor standard for de-identification and limited data sets. Any researcher using or sharing de-identified data obtained from a covered entity in a form that could possibly be used to re-identify an individual should evaluate whether the HIPAA de-identification standards are sufficient to protect their research and reputation.

NCVHS disagrees that HIPAA serves as the sufficient standard for de-identification of public health research covered by the Common Rule. Removing the set of 18 HIPAA direct identifiers is neither sufficient nor adequate to protect the identity of individuals in the current research environment. For example, outlying variable values and linkage of EHRs and other clinical data with public health program and surveillance data (or even with other publicly available data about individuals), potentially increase identification risk even after the HIPAA de-identification standard has been applied. NCVHS recommends developing more

²⁰ See Letters to the Secretary referenced in footnote 12

National Committee on Vital and Health Statistics

comprehensive approaches to the problem of unauthorized re-identification that will minimize these kinds of informational risks.

One specific recommendation might be to define the term ‘unauthorized re-identification’, explicitly prohibit unauthorized re-identification except in certain limited circumstances, establish penalties and other consequences for attempting to, or performing unauthorized re-identification, and create enforcement mechanisms to deal with unauthorized re-identification of information.

Moreover, it is important to ensure that researchers do not themselves attempt to re-identify information, either through use of a data set they are given or through combination of the data set with other sets of information. A data use agreement (DUA) is the current enforcement method for Limited Data Sets. The requirements for such a DUA include a provision prohibiting the recipient from identifying the information or contacting the individuals.²¹ This method relies on the supplier of the data gaining knowledge of any violations of the DUA and taking steps to enforce the contractual stipulations in the DUA. NCVHS is concerned about whether in the current economic environment where resources are scarce, diligent auditing of research partners is being carried out routinely. It is a significant challenge for data custodians merely to respond to the growing number of requests for data to ensure that a disclosure is within their data use policy, and we expect follow-up after the fact to be limited.

However, no DUA is required by HIPAA for de-identified information, and so no enforcement is available if a researcher or other party attempts to re-identify that information, even though the purpose of the de-identification was to prevent the possibility of identification.²²

As an additional problem NCVHS notes that the HIPAA standards for de-identification and for limited data sets apply to disclosure of data sets prior to their use for research purposes. They do not consider the subsequent question of whether research results, once released, can or will be re-identified. In the case of de-identified data, once it is disclosed, the recipient is not bound by the HIPAA Privacy Rule at all. Researchers reporting results in statistical form would, nevertheless, do well to evaluate the likelihood that individuals could be identified based on their results.

Therefore, these standards only go so far to protect subjects, and NCVHS has heard evidence that the field is evolving, that existing standards do not resolve all problems related to re-identification of individuals or small groups, and that other methodologies are being developed as techniques for re-identification grow more sophisticated. For example, some of the best practices are emerging due to the development of all payer claims databases and regulations associated with them at the federal and state level. Some of these projects use DUAs structured as a license that lasts one year, with a requirement to go back to the source if the recipient continues to require use of the data. This triggers the original custodian with an opportunity to do some kind of audit, investigation, or more informal follow up to ensure that their data policy is being followed. States could have regulatory follow up, audit, or affirmation that data have been returned or destroyed.

²¹ See 45 CFR §164.514(e)(4)(ii)(C)(5).

²²In contrast, many states require the execution of a DUA even when they give out a public use file which is supposed to be de-identified. However, state standards for de-identification may not be as stringent as the HIPAA Privacy Rule.

National Committee on Vital and Health Statistics

Another possible solution would be to have HIPAA covered entities be responsible for the errors of their research partners in the same way as they are for their business associates.

The goal in protecting research subjects is to establish better and more stringent guidelines, thresholds, guidance, and best practices (including statistical parameters for different scenarios of result releases) to ensure that the risk of unauthorized re-identification of study results is minimized or eliminated. And these guidelines and parameters must be periodically updated, as new, more advanced technologies for record linkages are developed.

In the case of an “excused” study, the ANPRM contemplates research that involves the use of data collected for other purposes, such as HIPAA de-identified data or limited data sets would not require consent, but identifiable data would require consent. These data would be subject to the new security requirements, but such a regime would not address the shortcomings we have discussed above.

An additional question posed here is whether the HIPAA structure is appropriate for social/behavioral research. This question is also posed in #59, and NCVHS addresses it below.

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?

The question has two parts: Would study subjects be adequately protected if researchers were subject to security standards based on the HIPAA Security Rule? And, second, are the HIPAA Security Rule standards appropriate for all kinds of studies?

NCVHS believes that the standards underlying the HIPAA Security Rule are a good basis for a security regime because these standards encompass the various facets of appropriate security practice: physical, technical, administrative, and policy controls, including risk assessments. They are only a beginning, however, for the research context. What is lacking across the board is a well documented understanding of how each of the Security measures (identity management, authentication, authorization, access control, audit, storage controls, controls of data in transit, etc.) applies to data in a research environment. One specific such application would be having a more comprehensive ‘module’ (component) within the security risk assessment and risk analysis tools that focuses on the collection, access, use and disclosure of data for research (in all its forms, including clinical, health services, applied, and public health research).

NCVHS also agrees with the need for physical, technical, and administrative standards for secure data storage, beyond the basic requirements already made part of the HIPAA Security regulations. These safeguards include standards for encryption of data at rest, stringent access controls of research data, and better security audit logs. In addition, researchers need to be more knowledgeable about security protocols for data storage and transmission used with research data. IRBs also should expand their review of technical, physical and administrative protocols for electronic data storage to strengthen general and public health research subject protection.

National Committee on Vital and Health Statistics

The HIPAA Security Rule extends security protections to all information it covers, some of which does not directly involve health. Some HIPAA-covered entities such as health plans receive and maintain personally identifiable information about individuals that is unrelated to their health status (e.g., social security numbers). Understanding what standards to apply to data protection must be grounded in a reasoned risk assessment able to classify the sensitivity of the information, identify the threats that exist to the privacy and security of the information, evaluate the potential harms, and apply the appropriate privacy and security measures that are commensurate with the need to protect the information and the individual from the harm. Such flexible assessments are the cornerstone of the HIPAA Security Rule and intended to assure that appropriate and adequate standards are considered.

It is often very difficult to isolate and identify all data elements and specify different requirements for different data types. As an example, the encryption of laptops is not mandated by HIPAA for PHI (although a risk assessment is required), but may be specifically required for personal information under more detailed state law. To determine what data a user may choose to download to a laptop and then determine if that laptop requires encryption is unworkable on a case-by-case basis. It is therefore more appropriate to require that all laptops used in an environment where identifiable information is handled are routinely encrypted in the event a user may, at some point, choose to store PHI or other personal information on the device. (Under the HIPAA Rules, a lost or stolen unencrypted laptop is assumed to contain PHI or PI unless it can be proven otherwise and therefore potentially to require notification of affected individuals).

Therefore, NCVHS recommends that if HHS proceeds down this path of considering security standards for all data in research, it should develop a more comprehensive information security risk assessment tool appropriate to research, and include guidance as to how each of the technical, physical and administrative safeguards are applied to data used for research. Extending the security requirements and standards under Meaningful Use (both Phase 1 and the upcoming Phase 2) to the use and disclosure of data for research also would be valuable. This would be part of the need to ‘harmonize’ the HIPAA Security requirements with those now required under Meaningful Use.

In considering adoption of security standards for data used in research, HHS should address the different phases of the life cycle of the data: 1) the collection by, or transfer of data to researchers; 2) the maintenance, use, or manipulation of data in the hands of researchers; 3) the release of data in public use files or as results in aggregate statistical form; and 4) the eventual disposition or destruction of the data. One aspect of security is the appropriate de-identification of the data to prevent a loss or misuse that could harm subjects, but, as we have said, the HIPAA Privacy Rule’s de-identification standards only apply to the first of these phases. Security standards must ensure that identifiable information does not leak during the researcher’s manipulation of data, during the release of research results, or during the wind-down, and disposal of data rely on other industry and statistical methodological standards, only some of which are the basis of the HIPAA Rules.

A novel issue now raised by the use of various types of data in research is the evolving concept of ‘distributed’ queries for identifying and conducting research (including population/public health research). This method epitomizes the opportunities, but, at the same time, the risks, of re-identification, patient matching, and other data linkages across organizations.

National Committee on Vital and Health Statistics

In response to the second question as to whether the HIPAA Security standards are appropriate to all types of research, NCVHS observes that these more comprehensive standards (including such things as encryption of data at rest or in transit) will be appropriate and applicable to all types of studies, but for certain social/behavioral type of research there might be a need to add additional requirements, including specific identified thresholds for the granularity of population analysis beyond which risks for re-identification of individuals increases. Also, there will need to be strong and strict constraints and limitations that apply to the new ‘distributed’ query methods for gathering information and conducting research.

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.”)?

NCVHS believes that additional data security and information protection standards should be considered. They include: 1) a comprehensive risk assessment/risk analysis for data to be used in research; 2) encryption of data at rest and in transit; 3) documented application of identification, authorization, authentication and access control requirements for access and use of data for research purposes; 4) research audit logs; 5) defined thresholds for use and disclosure of de-identified data to minimize or eliminate unauthorized re-identification. NIST’s Guide to Protecting the Confidentiality of Personally Identifiable Information would be a good model on which to base these new requirements

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?

As we have discussed above, if the recipient entity is not covered by HIPAA, the DUA forms the only link to enforcement between the entity disclosing information and the recipient researcher. Without it, no enforcement is possible. Even with a Data Use Agreement, data may be re-disclosed and re-identified by the recipient without the originator of the data discovering the problem.

If both parties are covered entities, then perhaps it would be acceptable for disclosure of PHI for research purposes between these entities to occur without a data use agreement, as long as the HIPAA protections for the data would continue to apply. Under the current HIPAA Privacy Rule, a DUA is required for disclosure of a limited data set, even if both parties are HIPAA covered entities.

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?

NCVHS believes that the ability to re-identify de-identified data is a very problematic gap in the current HIPAA Privacy Rule. Once data have been de-identified in accordance with the HIPAA Privacy Rule by removal of the 18 direct identifiers, the data no longer qualify as PHI and are completely removed from purview of HIPAA. The Office for Civil Rights has no authority over this data, as it is not covered by HIPAA. Since the entire purpose of de-identifying data is to remove the risk of identification, it would seem to undermine this purpose

National Committee on Vital and Health Statistics

if recipients could re-identify the data without penalty. While it is not a violation of the rule to do so, NCVHS regards such re-identification as an ethical violation by a researcher who attempts it or permits others to attempt it. Therefore, NCHVS recommends defining “unauthorized re-identification” and enforcement mechanisms to avoid the risks associated with this problem.

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

The risk of re-identification is at least the same for a trusted researcher as it is for an unknown third party who may receive data. Data that are collected by or disclosed to a researcher for a particular research project could lead to unfair surprise and loss of trust on the part of research subjects if, in the course of the research, the data are disclosed to a legitimate recipient who has not made the same promises regarding the protection against re-identification of the data. There should be some mechanism to protect these data from re-identification and potential misuse, including, potentially an absolute prohibition on re-identification by any party.