October 27, 2010

The Honorable John Berry
Director
Office of Personnel Management
1900 E Street, NW
Washington, DC 20415

Dear Director Berry:

We are writing to express our concerns about the Health Claims Data Warehouse (the "Warehouse") announced by the Office of Personnel Management (OPM) in the Federal Register on Oct. 5, 2010.[1] We have two major concerns:

(1) OPM's system of records notice (SORN) announcing the Warehouse appears to be legally deficient. The SORN does not allow for a fair public evaluation of the Warehouse or the role of OPM and the Dept. of Health and Human Services (HHS) in operating the Warehouse. We understand that a new SORN is being drafted and we applaud OPM for being responsive to the SORN's shortcomings. We outline below specific concerns with the initial SORN and urge you to address them in the revision. At minimum, the revised SORN should provide much more detail on how the data will be protected and used, and it should provide another, genuine opportunity for public comment. OPM should not create this massive database full of detailed individual health records without giving the public a full and fair chance to evaluate the specifics of the program. The Warehouse is planned to launch on Nov. 15, 2010. Since a revised SORN has yet to be issued, this date clearly needs to be delayed.

(2) OPM's SORN issues aside, there are more fundamental problems with the proposed Warehouse. Although OPM's stated goals for the Warehouse of reducing health care costs and boosting efficiency are laudable, the planned database is unnecessary and raises significant health data privacy and security issues. In light of the substantive issues we discuss below, we urge OPM to halt efforts to establish this database. OPM should instead consider using a query system that keeps the data with the record holders – such as the health plans.

Description of the Health Claims Data Warehouse

As described in the Federal Register, OPM's Warehouse would contain records on three nationwide health insurance programs.[2] The Warehouse would set up data feeds with the three programs to collect a broad range of personal information on enrollees in electronic format, including Social Security Number (SSN), spouses and children, employment, and health care coverage, procedures, diagnoses, and payments. The

[1] 75 Fed. Reg. 61532.
[2] The health insurance programs are the Federal Employee Health Benefit Program (FEHBP), the National Pre-Existing Condition Insurance Program (NPECIP), and the Multi-State Option Plan. FEHBP mostly covers federal employees, retirees, and their spouses, while the other two plans are not limited to federal employees.

Warehouse would retain records for seven years. The Warehouse's records would be retrievable through a combination of name and SSN.

OPM states that the Warehouse's purposes would include

- Disclosing enrollees' information to law enforcement agencies for prosecutions and investigations of possible violations of laws or regulations,
- Disclosing enrollees' information to Congress in response to congressional inquiry at the request of the enrollee,
- Disclosing enrollees' information to federal agencies, courts, and other parties during litigation or administrative proceedings in which the government is authorized to appear,
- Disclosing enrollees' information to researchers inside and outside the federal government,
- Analyzing enrollees' information to evaluate health care programs, and
- "Other purposes."

According to the Federal Register SORN, the Warehouse's data would be protected "in a secured database on a secured system" and that access to the database would be restricted to employees with the appropriate clearance and a need to know to perform their official duties. The notice includes a general statement that the data will be de-identified "in many instances." OPM provides no other information regarding security and privacy protections for the Warehouse's data.

OPM's Federal Register Notice Is Defective for Lack of Specificity

The Privacy Act requires Federal agencies seeking to establish a new system of records about individuals to publish a notice in the Federal Register containing several specific elements.[3] In its SORN Guide, OPM states that its policy is to describe the physical, technical and administrative safeguards taken to protect the records.[4] The SORN Guide also instructs OPM to describe all types of records the system would contain, to identify *each data element* maintained on an individual, and to collect nothing more than that which the SORN describes. Finally, the SORN Guide instructs OPM not to use collected information for any purpose other than that which the SORN describes.[5] OPM personnel are required to comply with the SORN Guide or face administrative penalties.[6]

OPM's SORN announcing the Warehouse fails to satisfy these SORN Guide requirements. OPM's SORN does not describe safeguards that are "sufficient to ensure the security and confidentiality of records and to protect against any anticipated threats or hazards."[7] Moreover, the SORN does not address the fact that the data at issue is health data, which is subject to unique and highly developed protections under HIPAA. For example, the SORN states that enrollee data will be de-identified "in many instances." The SORN does not explain what information will be stripped from the

[3] 5 USC 552a(e).
[4] U.S. Office of Personnel Management, *System of Records Notice (SORN) Guide*, Pg. 25, Apr. 2010, http://opm.gov/privacy/SORNGuide.pdf.
[5] Id., Pgs. 21-23.
[6] Id., Pg. 8.
[7] Id., Pg. 25.

records, whether the de-identification method meets HIPAA standards, or whether OPM will use "limited data sets."[8] The SORN does not reveal when the records will be de-identified, save that it will be de-identified "in many instances," nor is there any indication as to why OPM cannot de-identify the data for all specific uses. Likewise, OPM does not explain whether the Warehouse will be secured in compliance with HIPAA or even FISMA data security standards.[9] The SORN does not specify whether researchers or other third parties accessing the Warehouse's data will be subject to a business associate agreement, a certification of confidentiality, or a data use agreement.

OPM describes the personal identifying information to be kept in the system to include "Name, Social Security Number, Date of Birth, Gender, Phone number etc" [sic]. Likewise, OPM's SORN describes an open-ended use of the records: "To disclose to program and policy staff at OPM… for other purposes." Agencies must be specific about the categories of records in the system, as well as the specific uses of the records, and the use of "etc" and "other purposes" fails to meet the statutory requirement for disclosure.[10] Keeping these data elements open-ended makes it impossible to place limits on OPM's collection of personal identifying information and on the purposes for which OPM personnel will use the records, as required in the SORN Guide.

At minimum, OPM should issue a revised SORN that corrects these deficiencies, especially the lack of specifics regarding privacy and security protections. Without these details, the public and industry experts cannot fully and fairly evaluate the Warehouse program and the purpose of providing the public notice under the Privacy Act would be lost.

The Health Claims Data Warehouse Is Unnecessary

OPM does not need to create the Warehouse in order to accomplish the purposes described in the SORN. The government, researchers, and covered entities already possess the necessary authority to carry out the described uses for the Warehouse's data. Rather than duplicate sensitive enrollee information by copying it into the Warehouse, government agencies and researchers could access data already routinely collected in the ordinary course of business by the health plans participating in the affected insurance programs.

Health plans are already authorized to release medical information for research and statistical analysis, provided certain privacy protections are in place.[11] For example, persons enrolled in FEHBP receive their insurance benefits through private health plans. As the program administrator, OPM should have the authority to request that plans participating in FEHBP run data analyses to measure quality and cost effectiveness of the program. The plans could then provide OPM with aggregated information in response to these queries – rather than exposing identifiable raw data on individuals. OPM can thus achieve the goal of effective and efficient administration of FEHBP without moving sensitive individual information from its origin. OPM should explain why it

[8] 45 CFR 164.514.
[9] 45 CFR 164.306-316.
[10] 5 USC 552a(e)(4)(B).
[11] 45 CFR 164.512-514.

believes it must have physical possession of the enrollees' data to achieve its research goals.

With respect to the other routine uses for information in the Warehouse, existing law already provides a path for these uses to be accomplished using information collected by health plans in the ordinary course of business, subject to clear standards specified in HIPAA. For example, law enforcement already has the authority to obtain relevant health information from health plans when investigating or prosecuting violations of laws or regulations. Likewise, the government already possesses the authority to obtain relevant health information in litigation and administrative proceedings to which it is a party – again subject to rules specified in HIPAA.[12] In addition, many health plans already have internal units formed to investigate and report violations of the law, such as fraud, in partnership with law enforcement agencies.[13] There is no need to create a separate database, which, as far as the SORN is concerned, may or may not be subject to the HIPAA rules.

Congressional inquiries for an individual enrollee's medical information need not be directed at the Warehouse if – as the Federal Register SORN describes – the inquiry is made at the request of the enrollee. Such inquiries can also be directed at the health plans administering the programs in which the individual is enrolled. Individuals already have a legal right to obtain copies of their medical records from health plans in most circumstances.[14]

The Warehouse Would Exacerbate Privacy and Security Problems

The Federal Register SORN provides precious little detail regarding how OPM intends to secure the data in the Warehouse and protect individual privacy. Considering the wide range of information the Warehouse would collect, and the wide range of purposes for which OPM will use the data (some of which are open-ended), OPM should release specifics on the safeguards it will incorporate into the Warehouse. As it stands, the Warehouse appears to unnecessarily duplicate and centralize individual information, violate consumer privacy expectations, and exacerbate overuse of the SSN as an identifier – contrary to a White House directive and OPM's own policy.

In the SORN, OPM refers to the Warehouse as a "central and comprehensive database" that would contain large volumes of enrollee medical information that is also held in the enrollees' health plans. Duplication and centralization of data enhances the risk of security and privacy violations. Centralized databases are more prone to large-scale disruptions and breaches than decentralized models. Duplication of data increases the number of actors with access to the data, thereby making the data more vulnerable to misuse or unauthorized access. A decentralized – or federated – model is more resistant to these problems, can increase data quality, reduce costs, and provide greater value for

[12] 45 CFR 164.512(f).
[13] See, e.g., http://www.bcbs.com/blueresources/anti-fraud/what-the-blues-are-doing.html (last accessed Oct. 19, 2010).
[14] 45 CFR 164.524.

suppliers of health data.[15] In the context of the Warehouse, leaving individual records intact where they are housed or generated is a step closer to the decentralized database model.

As HHS has noted on numerous occasions, public trust is foundational to the health care system.[16] Yet the Warehouse runs contrary to individuals' expectations regarding what happens to their health data. Most people are aware (or should reasonably expect) that their health plans have medical and payment information from their claims. However, people are almost certainly unaware that the government intends to collect such detailed and sensitive information about them into one massive database in order to disclose the information to law enforcement and researchers, or for undefined "other purposes." In the case of the FEHBP, OPM currently assures enrollees that OPM does not have a copy of their personal health information.[17] Leaving individualized records with the health plans, subject to analyses and disclosure through the processes established under existing law, is more in line with the public's expectations of confidentiality.

OPM's SORN indicates that the Warehouse's record retrieval system will operate on a combination of individual names and SSNs. However, this would exacerbate the problem of overuse of the SSN in both government and the private sector. Overuse of the SSN has created significant identity theft concerns due to the SSN's central role in authenticating identity.[18] In 2007, as part of the President's Identity Theft Task Force, OMB issued a memorandum instructing government agencies to eliminate unnecessary collection and use of SSNs.[19] In 2010, OPM released a Privacy Impact Assessment which reiterated OPM's responsibility to reduce the use of the SSN.[20] Other federal agencies are in the midst of efforts to limit unnecessary use of the SSN.[21] Yet the Warehouse's record retrieval system would encourage health plans, researchers, and government agencies to link enrollee records to the SSN. The Warehouse's record retrieval system seems to be in direct conflict with federal policy to reduce SSN use.

The SORN states that enrollee data will be de-identified "in many instances." De-

---

[15] See Carol Diamond, Farzad Mostashari, and Clay Shirky, *Collecting And Sharing Data For Population Health: A New Paradigm*, Health Affairs, 28, no. 2 (2009): 454-466.

[16] See, e.g., U.S. Dept. of Health and Human Services, *Statement on Privacy and Security*, Jul. 8, 2010, http://healthit.hhs.gov/portal/server.pt?CommunityID=2994&spaceID=11&parentname=CommunityEditor&control=SetCommunity&parentid=9&in_hi_userid=11673&PageID=0&space=CommunityPage.

[17] U.S. Office of Personnel Management, *Frequently Asked Questions About Privacy of Medical Information*, http://www.opm.gov/insure/health/faq/privacy.asp#4 (last accessed Oct. 19, 2010).

[18] U.S. Federal Trade Commission, *Security in Numbers – SSNs and ID Theft*, Dec. 2008, http://ftc.gov/os/2008/12/P075414ssnreport.pdf.

[19] U.S. Office of Management and Budget, Memorandum 07-16, *Safeguarding Against and Responding to the Breach of Personally Identifiable Information*, May 22, 2007, http://www.whitehouse.gov/OMB/memoranda/fy2007/m07-16.pdf.

[20] U.S. Office of Personnel Management, *Privacy Impact Assessment Guide*, Pg. 3, Apr. 2010, http://opm.gov/privacy/PIAs/PIAGuide.pdf.

[21] See U.S. Dept. of Defense, Proposed Rule, *Reduction of Use of Social Security Numbers in the Dept. of Defense*, 75 Fed. Reg. 9548, Mar. 3, 2010. See also U.S. Dept. of Veterans Affairs, *Social Security Reduction Effort*, http://www.privacy.va.gov/ssn.asp (updated Mar. 23, 2010).

identification of data does not make the data risk free. Changes in society and technology have made re-identification of de-identified health information easier and cheaper than ever before, and the HIPAA Privacy Rule has never included mechanisms for holding recipients of de-identified data accountable for re-identification.[22] Moreover, health plans already have broad discretion to release de-identified data for a variety of purposes, so it is unclear why OPM must get involved in this business directly.[23]

Finally, the SORN contains blanket statements regarding how the Warehouse will disclose raw health data for research. The law requires covered entities using protected health information for research to first obtain individual authorization in most circumstances.[24] Yet OPM gives little indication regarding what types of research will be prohibited (e.g., are commercial uses permitted?), what criteria will be used to select research subjects, whether the results will be shared with the public, and to what extent individuals will have any choice in participating in the research. It would be irresponsible for OPM to move forward on this proposal without more clarity.

Consider Alternatives to the Health Claims Data Warehouse

The SORN does not adequately explain why OPM believes it must have physical possession of the enrollees' data. Since the Warehouse itself appears to be unnecessary to accomplish the purposes OPM lists in the Federal Register, the primary purpose of the Warehouse seems to be administrative convenience. That is, law enforcement, researchers, Congress, and other bodies would be able to exploit OPM's massive centralized database rather than having to request data from multiple health plans, as they can now. OPM should consider effective alternatives that would not violate the public's expectations of privacy or create unnecessary privacy and security problems.

An alternative would be to leave raw enrollee data with the current record holders – such as the health plans – and use a query system that can search diverse databases. The Food and Drug Administration already operates a similar system, called the Sentinel Initiative.[25] Sentinel was launched in 2008 in order to quickly monitor the safety of products the FDA regulates. Through Sentinel, the FDA can query product data and send questions to the data holders (which include health plans), but the data remains with and is managed by the participating data holders. Sentinel also operates under established privacy and security standards aimed at constant protection of personal information.[26]

Using a query system would be a closer fit with the public's expectations, leverage

---

[22] Center for Democracy & Technology, *Rethinking the Use of, and Rethinking Protections for De-identified (and "Anonymized") Health Data*, Pgs. 7-8, Jun. 2009, www.cdt.org/healthprivacy/20090625_deidentify.pdf.
[23] 45 CFR 164.502(d).
[24] 45 CFR 164.508.
[25] U.S. Food and Drug Administration, *FDA's Sentinel Initiative*, http://www.fda.gov/Safety/FDAsSentinelInitiative/default.htm (last updated Oct. 2010).
[26] U.S. Food and Drug Administration, *The Sentinel Initiative*, May 2008, Pg. 13. http://www.fda.gov/downloads/Safety/FDAsSentinelInitiative/UCM124701.pdf.

existing databases, minimize data transfer, and mitigate security risks. We urge OPM to consider a query system as an alternative to the Warehouse and to provide the public with substantially more detail regarding its proposal. Thank you.

These comments are submitted by the Center for Democracy & Technology (CDT) and the following additional supporters:

American Civil Liberties Union
American Federation of Government Employees, AFL-CIO
Childbirth Connection
Children's Partnership
Connecticut Health Policy Project
Consumer's Union
Domestic Violence & Mental Health Policy Initiative
National Alliance to End Sexual Violence
National Family Caregivers Association
National Partnership for Women & Families
National Network to End Domestic Violence
New Hampshire State Rep. Cindy Rosenwald, (D-Nashua)
New Hampshire State Rep. Neal Kurk, (R-Weare)
Patient Privacy Rights
Secure ID Coalition

For more information, please contact

Harley Geiger
Policy Counsel
Center for Democracy & Technology
202-637-9800 x 125
harley@cdt.org

Deven McGraw
Director, Health Privacy Project
Center for Democracy & Technology
202-637-9800 x 115
deven@cdt.org